

RESEARCH

Open Access



# Novel neural network classification of maternal fetal ultrasound planes through optimized feature selection

S. Rathika<sup>1</sup>, K. Mahendran<sup>2\*</sup>, H. Sudarsan<sup>3</sup> and S. Vijay Ananth<sup>4</sup>

## Abstract

Ultrasound (US) imaging is an essential diagnostic technique in prenatal care, enabling enhanced surveillance of fetal growth and development. Fetal ultrasonography standard planes are crucial for evaluating fetal development parameters and detecting abnormalities. Real-time imaging, low cost, non-invasiveness, and accessibility make US imaging indispensable in clinical practice. However, acquiring fetal US planes with correct fetal anatomical features is a difficult and time-consuming task, even for experienced sonographers. Medical imaging using AI shows promise for addressing current challenges. In response to this challenge, a Deep Learning (DL)-based automated categorization method for maternal fetal US planes are introduced to enhance detection efficiency and diagnosis accuracy. This paper presents a hybrid optimization technique for feature selection and introduces a novel Radial Basis Function Neural Network (RBFNN) for reliable maternal fetal US plane classification. A large dataset of maternal–fetal screening US images was collected from publicly available sources and categorized into six groups: the four fetal anatomical planes, the mother’s cervix, and an additional category. Feature extraction is performed using Gray-Level Co-occurrence Matrix (GLCM), and optimization methods such as Particle Swarm Optimization (PSO), Grey Wolf Optimization (GWO), and a hybrid Particle Swarm Optimization and Grey Wolf Optimization (PSOGWO) approach are utilized to select the most relevant features. The optimized features from each algorithm are then input into both conventional and proposed DL models. Experimental results indicate that the proposed approach surpasses conventional DL models in performance. Furthermore, the proposed model is evaluated against previously published models, showcasing its superior classification accuracy. In conclusion, our proposed approach provides a solid foundation for automating the classification of fetal US planes, leveraging optimization and DL techniques to enhance prenatal diagnosis and care.

**Keywords** Ultrasound Images, Fetal organs, Fast Radial Basis Function Neural Network, Optimization, Features, Accuracy, Data Balancing

## Introduction

US imaging is a safe and effective method for monitoring the health of the baby and mother during pregnancy [1, 2]. It is non-invasive, offers real-time imaging capabilities, and is reasonably priced [3]. Commonly used during obstetric examinations, it helps achieve the best possible outcomes for both the mother and child. The most typical timing for a prenatal US is during the third trimester, typically between weeks 18 and 24 [4]. Obstetricians can

\*Correspondence:

K. Mahendran  
srimahendrancs@gmail.com

<sup>1</sup> Prince Shri Venkateshwara Padmavathy Engineering College, Chennai, India

<sup>2</sup> Saveetha Engineering College, Chennai, India

<sup>3</sup> K. Ramakrishnan College of Engineering, Trichy, India

<sup>4</sup> Chennai Institute of Technology, Chennai, India



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

measure and track fetal growth and development by capturing normal fetal planes during routine clinical examinations. Most countries provide prenatal care, which includes at least one scan in the third trimester to detect any anomalies. Fetal weight measurement is crucial for diagnosing anomalies and monitoring growth. Obstetricians calculate fetal weight using standard fetal US planes, which include four biometric parameters: abdominal circumference, femur length, biparietal diameter, and head circumference [5]. Maternal–fetal US images are frequently used to assess the health of the unborn child by evaluating biometric data such as gestational age and weight. The majority of fetal medicine institutes adhere to international standards established by scientific organizations when obtaining US images of both the mother and the unborn child. This ensures reproducible image acquisition using standardized approaches. Screening US performed in the middle trimester frequently provides more than 20 images per session. To enhance clinical evaluation, three-dimensional (3D) images, and films are sometimes acquired [6]. Fetal specialists select images that include structures of interest after analyzing the sonographer's findings. Confirmation by a senior maternal–fetal expert follows a team of certified research technicians' initial assessment. However, this procedure is challenging, error-prone, and time-consuming due to the large number of images in each screening US scan [7, 8]. Thus, a low-cost, presumably error-free automated system could perform the job. World Health Organisation (WHO) data indicate that sonographers in developing countries often lack the necessary education to conduct fetal US scans [9]. To overcome these challenges, automated classification of normal maternal–fetal US images can be implemented. This will aid inexperienced sonographers, reduce the burden of obstetricians, and increase diagnostic efficiency.

The recent decade has witnessed tremendous improvements in Artificial Intelligence (AI), particularly with the introduction of DL, and its success in image recognition tasks [10, 11]. DL has proven effective in various medical applications, including classifying and segmenting organs and lesions in computer tomography and magnetic resonance imaging images. These approaches excel at automatically detecting complex patterns in picture data and providing quantitative and qualitative assessments [12]. Early diagnosis of abnormalities relies on rapid and accurate assessment of US images to ensure the safety of both mother and fetus. This study aims to advance the field by presenting a hybrid optimized technique for better feature selection in maternal–fetal US images, as well as a fast RBFNN for accurate classification in this domain.

Gynecologists utilize US images to detect pregnancy problems and monitor the progress of a developing fetus.

US is widely employed in the medical field for categorizing fetal standard planes. This article describes a unique methodology for detecting maternal–fetal planes in US images based on the RBFNN architecture. RBFNNs are a popular type of neural network used in computer vision and image processing. The primary contributions of this study are as follows:

1. Maternal fetal US images were collected from a large publicly accessible dataset.
2. Several preprocessing techniques are implemented, such as scaling, removing undesirable portions from US images, improving image quality, and extracting features.
3. Use the hybrid optimized PSOGWO method to choose the most significant features, enhancing the classifier's accuracy.
4. The novel DL model, Fast-RBFNN is proposed with  $\epsilon$ -insensitive loss function and structural risk term. This technique not only maintains robust nonlinear fitting with easy learning rules, but it also handles huge data sets quickly and efficiently.
5. Evaluate the proposed DL model against conventional DL models such as Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and Radial Basis Function (RBF). Additionally, a comparison with current works is performed.

The study's framework is presented below. The "Related work" section includes an overview of research associated with this domain. The "Materials and Methods" section details data collection, processing, feature extraction, selection, and classifiers. The outcome of the proposed model is discussed in the "Results and Discussion" section. Finally, the "Conclusion" section discusses the effectiveness of the proposed methodology and potential areas for future research.

## Literature survey

AI systems' predicting and classifying skills are driving wider applications in clinical research. As a result, they have found widespread application in biomedical research and the development of reliable tools for diagnosis. Numerous studies have been conducted to classify maternal–fetal standard planes using ultrasonography.

Image categorization, facial recognition, and speech recognition are just a handful of the challenges that AI has helped overcome. However, this strategy will not be very effective for classifying and segmenting US images. The automatic categorization of anatomic planes from fetal US images is complicated by several issues, including poor SNR, the mother's bulk, and the small size of the fetus [13]. The problems are made worse by the fact

that images in hospital archives are taken with a range of equipment types, operators, imaging formats, resolutions, zoom settings, and so forth. Various DL models like GoogLeNet [14] and ResNet [15] have been proposed to overcome the above challenges. The use of these algorithms in classification tasks involving medical images yields excellent results. The study [16] used Deep Convolutional Neural Networks (DCNNs) to increase recognition accuracy and optimize clinical pipelines by identifying fetal facial standard planes (FFSPs). Training DCNN models with improperly labeled examples caused overfitting and performance loss. The study [17] proposed a Transfer Learning (TL) strategy to overcome the issues of limited training data and diminishing performance. This method involves implanting the knowledge of specific CNN models that have been pre-trained on a broad scale with images of natural settings. These models are used to localize medical images such as fetal abdominal standard planes. When it came to the auxiliary task of plane recognition, the metrics measurements demonstrated that DCNN and TL were effective. The algorithm's limitations in real-time manipulation and reliance on quantitative medical datasets rendered it unsuitable for practical usage. Recent works [18] propose a unique 82-layer DL architecture based on residual bottleneck mechanisms. The intended design now includes three more blocks, each with highway routes and skip connections. Furthermore, before each residual block, a convolutional layer has been added. During training, several hyperparameters were initialized using Bayesian optimization (BO). The classification is performed with deep features taken from the average pooling layer. A more efficient Moth Flame Optimization strategy for feature selection was proposed to address increased computation time during the classification phase. Next, the selected features are used to train neural network classifiers to categorize the data. During the experimental phase, US images were evaluated, with a focus on normal mother-fetal images and images of the fetus's developing brain. Some limitations are (i) A bias in the dataset complicates DL model training; and (ii) the deep layer extracts irrelevant data.

The research [19] aims to improve fetal brain US plane categorization by investigating and evaluating the use of a Generative Adversarial Network (GAN) to generate synthetic US images of the brain. The synthesis of fetal brain images utilizing cutting-edge GANs stylegan2-ada was compared to baseline classifiers that used GAN-based data augmentation. Our experiments suggest that combining GAN-generated data with standard augmentation techniques improves accuracy. The focus was solely on binary categorization, ignoring all other possible options. Using a brain detector, the US images of the developing

brain were pre-processed and centered. No evaluation was conducted on these architectures' performance on raw US images. The paper [20] proposed a novel framework called FetalBrainAwareNet for generating anatomically accurate synthetic images of fetal head standard planes (FHSPs). This framework enhances the presence of desired anatomical features in the generated images. It also explores specific regularization factors in the adversarial training loss function to control the fetal skull shape and improve the distinction of standard planes, ensuring that the synthetic images structurally and visually match real US scans. The framework demonstrates its versatility by generating high-quality images of the three most common FHSPs. Quantitative and qualitative results indicate that the system produces a greater variety of US images compared to existing methodologies. In the article [21], a new U-Net-based network for medical image segmentation was developed. This enabled us to approach the problem. The architecture of the proposed model consists of four layers: two-dimensional convolutional, two-dimensional transposed convolutional, and batch normalization. The encoder-decoder path consists of four blocks. The performance of the proposed network was tested on a publicly available database for fetal head circumference estimation and a newly created database for head and belly circumference. The suggested fast network significantly reduced processing time when compared to existing U-Nets. The suggested model improved segmentation accuracy since it has more trainable parameters than comparable models. The researchers [22] present a method for automatically identifying 14 distinct fetal characteristics in 2-dimensional fetal US images by integrating data from the entire image and selected areas of interest (AOI). Our solution uses pre-trained CNNs to learn two feature extractors on entire as well as AOI of US fetal images. Our approach is unique in that it combines classification decisions based on global and local features without the usage of prior knowledge. Furthermore, our method can identify fetal features in the image based on the categorization results. In comparison to the other non-fusion-based techniques, the proposed solution is statistically significantly better. Because of this limitation, the developed AI algorithms may be less effective or generalizable, and they will be unable to handle a broader range of clinical circumstances.

To improve the accuracy of classical ultrasonic plane detection, graph-enhanced multi-scale structure perception architectures were used [23]. Specifically, a graph-based multi-view refinement module was introduced for detecting linkages between fetal anatomical components, along with a local-to-global multi-granularity ensemble module for feature enhancement and noise suppression. A confidence assessment loss was added, and several

classifiers with branch subnets were employed to produce fine-grained structure representations. Confidence matching enables classifiers to collaborate while generating choices. The paper [24] proposes an automated method for recognizing common fetal US planes using a CNN stacking ensemble. This strategy employs three CNNs: AlexNet, VGG-19, and DarkNet-19. Predictions from these CNNs are extracted using random forest and softmax classifiers. The final forecast is determined using the absolute majority voting method. The stacking ensemble method was evaluated on a freely available prenatal US dataset, showing superior performance compared to competing approaches and individual CNN models. The study [25] explores the effectiveness of SimCLR in scenarios with low and high inter-class variability, considering that classification performance varies with the number of labels. Various training approaches utilizing contrastive learning were employed, alongside quantitative and qualitative research using industry-standard measurements. Contrastive learning proved to be more effective for low inter-class variability classification tasks, particularly when trained with ImageNet weights, compared to high inter-class variability tasks.

The study [13] proposes an automated approach for identifying regularly utilized fetal US planes using the DCNN stacking ensemble. All three DCNNs have been pre-trained using the stacking ensemble technique. Random forest and softmax classification algorithms are utilized to make predictions for DCNNs. The final forecast is generated with the majority vote method. The stacking ensemble method is evaluated on a freely available fetal US dataset. The suggested ensemble model for fetal US planes establishes six categories in total. The experiment results reveal that the stacking ensemble method was extremely effective. Training DCNN models with improperly labeled examples caused overfitting and performance loss. Article [24] describes a DL-based automated categorization system for fetal US planes to improve detection efficiency and accuracy. The suggested solution is mostly made up of feature integration and categorization modules. Following the retrieval of deep features using TL models, the global average pooling layer is first merged. Merging deep features from multiple CNNs produces a more robust feature representation. The next stage is to employ a multi-layer perceptron with deep features to predict fetal US images. When compared to other cutting-edge models that are already available, the proposed technique demonstrated higher categorization efficiency. This method's possible downsides include difficulty with real-time use, overfitting, model generalization, and dataset quality [26].

In existing research, many DL models have been developed, but they still struggle to achieve excellent accuracy in classifying fetal organs from US images. These models can be adversely affected by noise present in medical images, and some fail to extract important features from US images effectively. Feature extraction is crucial for obtaining better results. Additionally, some models suffer from overfitting problems. To address these issues, this research presents a hybrid PSOGWO technique for effective feature selection and proposes Fast-RBF with an  $\epsilon$ -ILF and structural risk minimization to avoid overfitting.

### Proposed methodology

Publicly accessible US image data containing maternal–fetal US planes are gathered for analysis. Raw US images undergo careful pre-processing to enhance their quality and suitability for classification purposes. The pre-processing steps are detailed in Section III.B. GLCM is employed to extract features, that are effective for texture analysis—a crucial aspect in US imaging where texture patterns provide significant insights into the anatomical structures of fetal planes. GLCM extracts 45 distinct features from each image, capturing essential texture information for classification. To improve classification accuracy and speed, optimization techniques are utilized to reduce the feature dimension. The PSO algorithm is well-known for its ability to solve a variety of real-world issues. However, PSO might get caught in local minima. To address this problem, the GWO is employed to assist PSO, preventing local minima and improving the search process. PSO, GWO, and their hybrid, PSOGWO, are used to identify the most important features. Through this process, the original set of 45 features is reduced to 25.

The selected features from each optimization algorithm are inputted into various DL models, including ANN, CNN, RBF, and the proposed DL model. The performance of each combination of the feature selection algorithm and DL model is evaluated using appropriate metrics. The model achieving the highest classification performance based on the evaluation metrics is identified as the best classifier for maternal fetal US planes. The proposed methodology for maternal fetal classification is visually represented in Fig. 1.

### Data acquisition

The fetal dataset from Zenodo [27] was used for this study. The collection consists of 12,400 US pictures. Images were divided into five anatomical planes (brain, abdomen, femur, thorax, and maternal cervix) that are

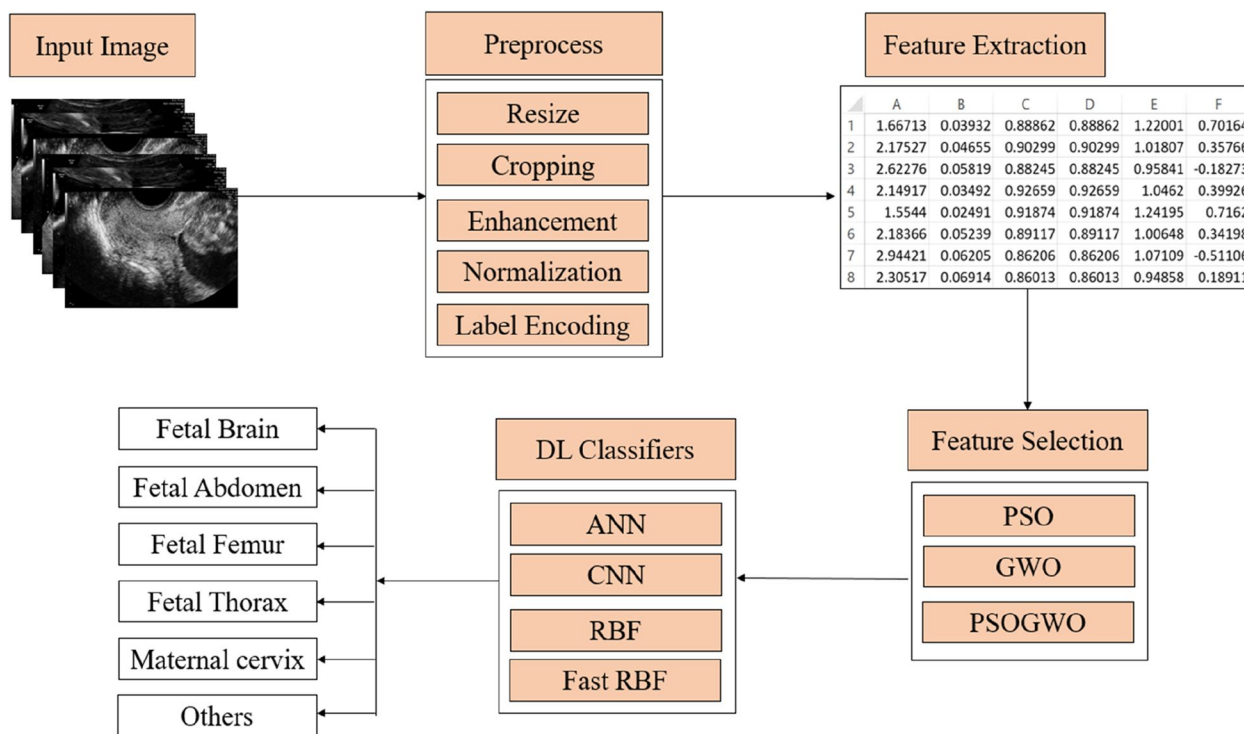


Fig. 1 Proposed methodology workflow

most commonly used for maternal–fetal screening, as well as an "Other" category that includes additional planes. Figure 2 shows samples of US images from each category. The details of images are listed in Table 2. According to this table, the number of images in each type ranges from 714 to 4,213. As a result, it appears that there is an issue with class imbalance.

**Data processing**

To learn image-based AI models, images need to be first processed. In the absence of pre-processing, Low-quality input data might cause AI models to underperform. Several methods are employed in pre-processing techniques to improve the input data, such as resizing, cropping, noise removal, image enhancement, balancing, and encoding.

Variations in image size in raw US images can impair model accuracy [28]. Images are scaled down to a standard resolution while maintaining their aspect ratio to ensure consistency and save computing load. The image resize process can be described in Eq. (1):

$$Resized\ Image = resize(Original\ Image, Target\ Resolution) \quad (1)$$

where the raw US image is referred to as the *OriginalImage*. The *TargetResolution* specifies the ideal image resolution to be achieved when resizing.

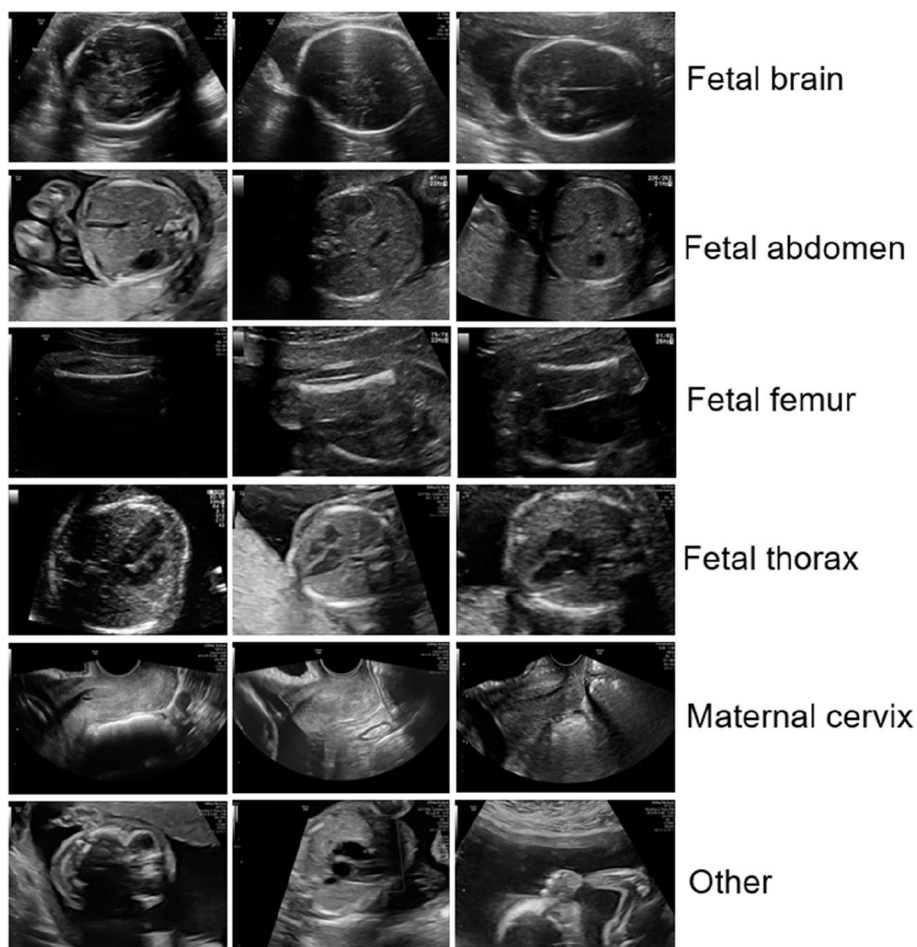
Cropping is a typical technique for focusing on a specific part of an image and removing distracting background components [29]. There is a lot of irrelevant information (patient data, system settings, image position, etc.) in the provided dataset that ruins the images’ ability to show the AOI. Model reliability and image readability are improved by cropping. It is a simple and effective strategy for improving image quality, resulting in more relevant and legible data for the suggested models. By reducing superfluous data, our cropping method improves model accuracy and speeds up training. When cropping an image, the height (*h*) and width (*w*) of the edges are first measured. The cropping borders are then set, ensuring that the *x* and *y* values remain within permissible limits. The cropping method uses the computed *x* and *y* values, as well as *h* and *w* values.

Image processing techniques are used to improve the contrast of US images and uncover previously concealed elements. The term "histogram equalization," is in Eq. (2):

$$Enhanced\ Image = histeq(Cropped\ Image) \quad (2)$$

Here, the image after cropping is termed as *CroppedImage*. The histogram equalization operation is represented by *histeq*.

Normalizing pixel values to a constant range is essential for reliable model training. Normalizing pixel values



**Fig. 2** Sample US images from the Zenodo dataset























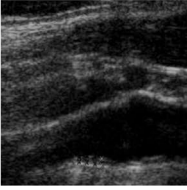
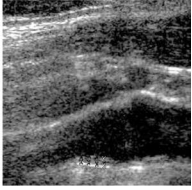
to the  $[0, 1]$  range is a common practice in image processing [30]. Each image has an individual number that corresponds to the name of the anatomical plane associated with it. Since most ML approaches require numerical inputs. The outcome of the pre-processed US image at each stage is given in Table 1.

Data augmentation approaches were employed to address class imbalance in the dataset's minority classes. Oversampling and undersampling techniques [31] were used to overcome this issue, as shown in Table 2. This strategy involved boosting and reducing the representation of minority and majority classes by randomly removing some cases and duplicating existing samples. Each group has been allotted an instance of 1050. This strategy, which distributes the dataset's classes more evenly, can improve the model's accuracy. Figure 3 depicts the US image distribution before and after applying the sampling approaches. This method significantly improves the effectiveness of the proposed models by rectifying the class imbalance [32].

#### Feature extraction

The texture is created by spatially alternating grayscale values, resulting in a spatial link between two pixels in the image separated by a set distance. An accurate approach to texture description is examining the spatial connection of grayscale values. The GLCM is a popular technique for showing the spatial correlation of pixel grayscale [33]. It characterizes a picture by taking into account the distance between adjacent pixels, the magnitude of variation, and the direction of correlation. The essential notion is to compute the frequency of occurrence of two grayscale pixels in a certain spatial arrangement, which can be used to assess the image's regional relativity and consistency. As the distance changes, GLCM evolves immediately in fine textures but slowly in coarse textures. The GLCM defines a square matrix whose magnitude corresponds to the likelihood that a gray value  $g_1$  is separated from another gray value  $g_2$  by a fixed spatial location connection (dimensions and orientation). Consider  $f(i, j)$  as a two-dimensional grayscale image, where

**Table 1** Pre-processed image outcome at each stage

	Actual Image	Resize Image	Crop Image	Enhanced Image
Brain				
Abdomen				
Femur				
Thorax				
Maternal Cervix				
Other				

$S$  is the collection of pixels within the region that have a certain spatial relationship and  $P$  represents the GLCM, which is expressed in Eq. (3)

$$P(i, j) = \frac{\#\{[(i_1, j_1), (i_2, j_2)] \in S | f(i_1, j_1) = g_1 \& f(i_2, j_2) = g_2\}}{\#S} \quad (3)$$

There are two techniques to obtain pixel information from GLCM: distance and angle orientation [34]. When the distance is too small, the data in each pixel becomes extremely constant. However, going overboard renders the data between pixels worthless. The distance  $d=5$ , and the scenario encompasses four angles:  $\{0^\circ, 45^\circ, 90^\circ \text{ and } 135^\circ\}$ .

**Table 2** US fetal image distribution

Classification	Actual	After balancing
Fetal Brain	3092	1050
Fetal Abdomen	711	1050
Fetal Femur	1040	1050
Fetal Thorax	1718	1050
Maternal Cervix	1626	1050
Other	4213	1050

The scenario applies to all patch images due to the presence of multipatch images. Feature extraction using GLCM retrieves data associated with specific variables, including energy, correlation, dissimilarity, homogeneity, and contrast. After converting RGB images to grayscale, GLCM extracts pixel information. GLCM utilizes two orientations: angle and distance. The co-occurrence matrix will be updated before the GLCM variable is computed. Divide the total number of recorded co-occurrences by the specified normalization. As previously stated, the variables will be calculated after normalization. This is a reference to the Haralick publications, which were critical to the development of GLCM [35]. Each of the variables is defined in Eq. (4–13).

$$Contrast = \sum_{i,j=0}^{N_g-1} p(i,j)(i-j)^2 \tag{4}$$

$$Dissimilarity = \sum_{i,j=0}^{N_g-1} p(i,j)|i-j| \tag{5}$$

$$Homogeneity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i,j)}{1+(i-j)^2} \tag{6}$$

$$Energy = \sum_i \sum_j \{p(i,j)\}^2 \tag{7}$$

$$ASM = \sum_i \sum_j \{p(i,j)\}^4 \tag{8}$$

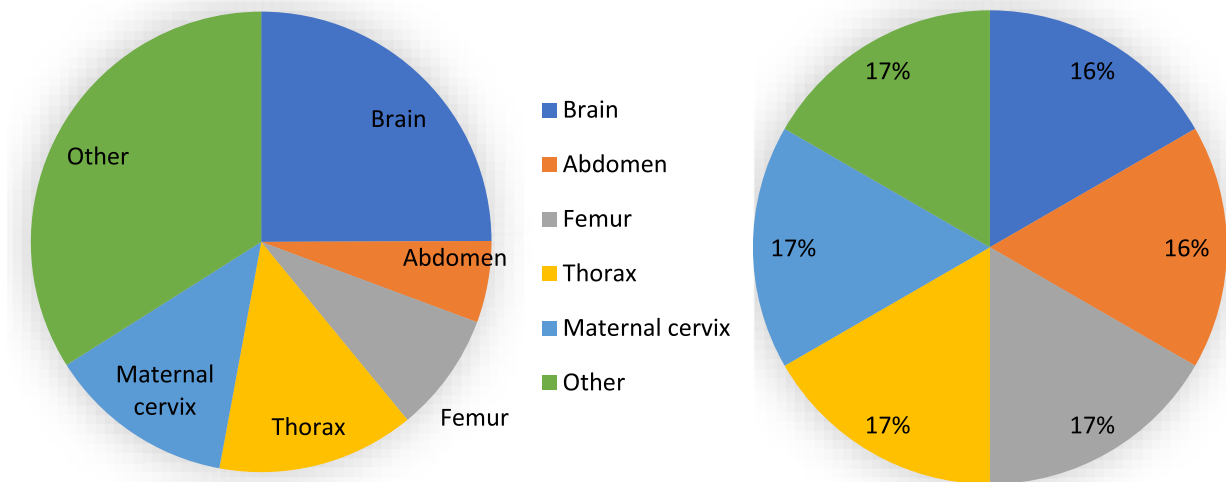
$$Correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} [(ij)p(i,j)] - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{9}$$

$$Entropy = - \sum_i \sum_j P(i,j) \lg P(i,j) \tag{10}$$

$$Mean = \mu_i = \sum_{i,j=0}^{N-1} i(P_{i,j}), \mu_j = \sum_{i,j=0}^{N-1} j(P_{i,j}) \tag{11}$$

$$Variance = \sigma_i^2 = \sum_{i,j=0}^{N-1} P_{i,j}(i - \mu_i)^2, \sigma_j^2 = \sum_{i,j=0}^{N-1} P_{i,j}(j - \mu_j)^2 \tag{12}$$

$$StandardDeviation = \sigma_i = \sqrt{\sigma_i^2}, \sigma_j = \sqrt{\sigma_j^2} \tag{13}$$



**Fig. 3** Distribution percentage of US images before and after sampling



The GLCM co-occurrence matrix element  $p(i,j)$  is normalized and symmetric in relation to  $(i,j)^{th}$ . Contrast is defined as the intensity of the pixels surrounding the reference pixel at a given angle and distance. In general, great contrast indicates a lively image. Dissimilarity will calculate the spacing between any two items in the area of interest, as represented by pixels. A higher value implies that there is a considerable difference in intensity between adjacent pixels. The homogeneity of a pixel's distribution in a GLCM is measured. The relationship between contrast and value will be inverse. The contrast decreases as the degree of homogeneity increases. ASM measures the regularity of pixels in GLCM. The higher the ASM number, the more similar the pixels are. ASM provides 100% of GLCM's power. ASM's base is energy. The correlation demonstrates a local gray-level dependency on the texture picture through the linear dependency of the gray-level value in the GLCM. A comparable gray-level area can produce stronger correlation results. The ultimate purpose of all of the equations is to supply six texture features per image patch to train the deep neural network. All four of these angles ( $0^\circ, 30^\circ, 45^\circ, 90^\circ$  and  $135^\circ$ ) represent the location of the focus pixels. Meanwhile, the distance between the reference and nearby pixels is  $d = 5$ . The position is determined first, and then each variable is calculated. A total of 45 features were collected from GLCM. The hyperparameters used in this study are detailed in Table 3.

**Feature selection**

**PSO**

In the research [36], the PSO algorithm was used for feature selection in COVID detection and achieved better results. Based on this, we decided to choose PSO. The algorithm 1 demonstrates how PSO represents the evolution of understanding of social behavior and the dynamics of group communication during the exchange of secret information regarding migration, flocking, or hunting [37]. Together, they form a solution; the former is called a swarm, while the latter is called particles. Particles can modify their positions by utilizing both their own and the information of their neighbors [38].

**Table 3** GLCM Hyperparameters

Hyperparameter	Values
Patches	1000
Distance	5
Angles	$0^\circ, 30^\circ, 45^\circ, 90^\circ$ and $135^\circ$
Number of Features	45

```

Algorithm 1. Pseudo code for PSO
Input: GLCM Features
Output: Optimal features
Randomly initialize population()
Randomly initialize position()
Randomly initialize velocity()
While termination criteria meet:
    Evaluate objective function (OF) using selected features
    Compute OF
     $ANN's\ Accuracy = \frac{Correct\ prediction}{Total\ prediction}$ 
    Calculate current position and velocity
    If current OF > global OF
        Optimal features = current selected features
    Else
        No change in Optimal features
    For each particle:
        Update position and velocity using Equation (14) and (15)
    End
End
Return optimal features
    
```

To begin, the swarm generates a set of random particles dependent on positions and velocities. The procedure for updating the particles' locations and velocity is represented by Eqs. (14) and (15):

$$x_{ij}^{(t+1)} = x_{ij}^{(t)} + v_{ij}^{(t+1)} \tag{14}$$

$$v_{ij}^{(t+1)} = wv_{ij}^{(t)} + c1r1(x_{ij}^{p(t)} - x_{ij}^{(t)}) + c2r2(x_{ij}^{g(t)} - x_{ij}^{(t)}) \tag{15}$$

Here,  $t$  represents the current iteration and  $w$  is an inertia weight, used to expedite population convergence. When the position of the  $i$ -th particle in the  $j$ -th dimension is denoted by  $x_{ij}$  and its velocity by  $v_{ij}$ . Additionally,  $c1$  and  $c2$  are acceleration coefficients, constant values. The previous best position of particle  $i$  in the  $j$ -th dimension is represented by  $x_{ij}^{p(t)}$  and  $x_{ij}^{g(t)}$  respectively. The parameters  $r1$  and  $r2$  can take values between 0 and 1. Then, the main loop of PSO evaluates each particle using a fitness function and compares the results with both local and global best values.

**GWO**

The GWO algorithm simulates the leadership structure and hunting strategies of grey wolves [39]. The alpha ( $\alpha$ ), beta ( $\beta$ ), delta ( $\delta$ ), and omega ( $\Omega$ ) packs of grey wolves serve as a model for the organizational structure. Additionally, the algorithm incorporates the three main components of hunting: seeking out prey, encircling prey, and attacking prey [40]. Algorithm 2 presents the pseudocode for the GWO algorithm.

```

Algorithm 2. Pseudo code for GWO
Input: GLCM Features
Output: Optimal features
Randomly initialize population()
Randomly initialize parameters (a, A, c)
While termination criteria meet:
    Evaluate objective function (OF) using selected features
    Compute OF
     $ANN's\ Accuracy = \frac{Correct\ prediction}{Total\ prediction}$ 
    Calculate best search agents ( $x_\alpha, x_\beta, x_\delta$ )
    If current OF > global OF
        Optimal features = current selected features
    Else
        No change in Optimal features
    For search agent:
        Update search agent's position using Equation (18)
    End
End
Return optimal features
    
```

$$D_\alpha = |\vec{c}_1 \cdot \vec{x}_\alpha - \vec{x}|, D_\beta = |\vec{c}_2 \cdot \vec{x}_\beta - \vec{x}|, D_\delta = |\vec{c}_3 \cdot \vec{x}_\delta - \vec{x}| \tag{16}$$

$$\vec{x}_1 = \vec{x}_\alpha - A_1 \cdot D_\alpha, \vec{x}_2 = \vec{x}_\beta - A_2 \cdot D_\beta, \vec{x}_3 = \vec{x}_\delta - A_3 \cdot D_3 \tag{17}$$

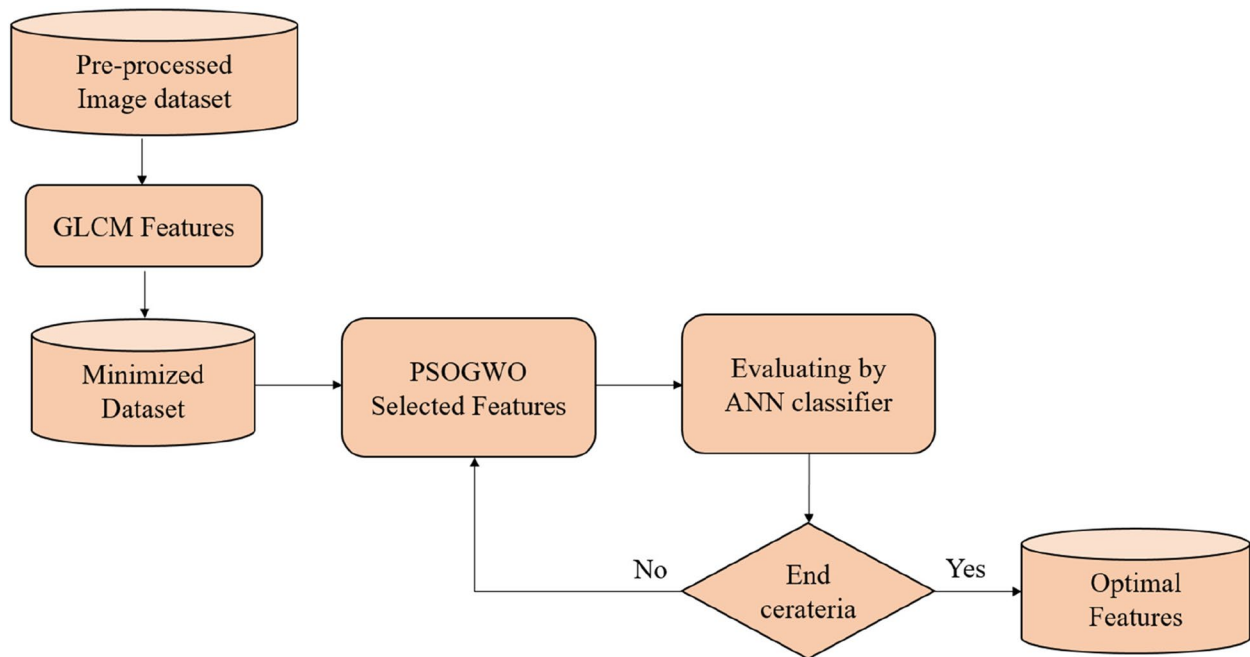
$$\vec{x}(t + 1) = \frac{\vec{x}_1 + \vec{x}_2 + \vec{x}_3}{3} \tag{18}$$

The distances between each  $\delta, \beta,$  and  $\alpha$  and the prey are represented by  $\vec{x}_1, \vec{x}_2,$  and  $\vec{x}_3$ .

**PSOGWO**

Figure 4 depicts the overall structure of PSOGWO for optimized feature selection. PSOGWO combines PSO and GWO wrapping algorithms. At this point, testing the effectiveness of the two algorithms, GWO and PSO, which employ different search methodologies, is desired. When it comes to feature selection algorithms, PSO is a favorite among researchers. The PSO's update mechanism is integrated into the fundamental design of GWO. The first two stages of PSOGWO are parameterization and population formation, which cover a wide range of problem solutions (feature selection). Calculating the fitness function for each solution and selecting the best one helps us to assess their utility. The next stage of the PSOGWO algorithm involves updating the population with the concurrent GWO and PSO algorithms. Following that, the grey wolf fitness function is compared to the best in the world, and the variables are adjusted as

Here is a brief rundown of the grey wolf hunting method: The  $\alpha, \beta,$  and  $\delta$  are likely the most knowledgeable about the potential whereabouts of prey. Alpha represents the best candidate solution. By retaining the top three results, it can be ensured that all search agents, including the omegas, will be compelled to adjust their positions to match those of the top search agents. Grey wolves use these formulas from Eq. (16–18) to keep track of their positions.



**Fig. 4** Working of PSOGWO algorithm for feature selection

needed. Equation (19) is utilized to determine the fitness function.

$$fitness = weight_{acc} * accuracy(agent) + weight_{feature} * \frac{tot\_feat - sel\_feat}{tot\_feat} \tag{19}$$

where *tot\_feat* is the total features in the agent, *sel\_feat* is the features selected by the agent, and *accuracy(agent)* is the agent’s classification accuracy. The parameters are then revised to reflect the new positions, and the grey wolves and global best adjust to match. Once the final requirements are met, the process is repeated. The ultimate result is a vector of integers indicating whether characteristics were selected or removed. As illustrated in Algorithm 3, the PSOGWO technique is recommended.

```

Algorithm 3. Pseudo code for PSOGWO
Input: GLCM Features
Output: Optimal features
Randomly initialize swarm population()
Randomly initialize grey wolf population ()
Randomly initialize velocity and position()
Randomly initialize parameters (α, A, c)
While termination criteria meet:
    Evaluate objective function (OF) using selected features
    Compute OF
    ANN's Accuracy =  $\frac{Correct\ prediction}{Total\ prediction}$ 
    Calculate current position and velocity
    If current OF > global OF
        Optimal features = current selected features
    Else
        No change in Optimal features
    For each particle:
        Update position and velocity using Equation (14) and (15)
    End
    For search agent:
        Update search agent's position using Equation (18)
    End
End
Return optimal features
    
```

**Classifiers**

Figure 5 depicts the three layers that comprise RBFNN: Input Layer (IL), Hidden Layer (HL), and Output Layer (OL) [41]. This includes  $x_i \in R^d$  and  $y \in R$ , with  $M$  hidden layer nodes, and the RBFNN executes the nonlinear mapping  $f : R^d \rightarrow R$ . The training US images are fed into the RBFNN’s input layer, while the HL nodes conduct a nonlinear transformation using the RBF for mapping the input and new space. When constructing a Gaussian function, the variables in the RBF reflect the centre ( $c_i \in R^d$ ) and kernel width ( $\delta_i$ ). The function is described in the Eq. (20):

$$\varphi(\|x - c_i\|) = exp\left(-\frac{\|x - c_i\|^2}{\delta_i}\right) \tag{20}$$

A linear weighted combination is applied in the new space by the OL’s nodes. For the mapping function from  $R^d \rightarrow R$ , let’s assume  $\varphi(\blacksquare)$  is the RBF and  $w_i$  is the connection weight between the HL and OL.

$$y = f(x) = \sum_{i=1}^M w_i \varphi(\|x - c_i\|) \tag{21}$$

With the use of a linear model and an RBFNN. As stated in the introduction, three parameters are needed for the RBFNN. The parameters are the centre vector of the RBFNN  $c_i = [c_{i1}, c_{i2}, \dots, c_{id}]^T$ , the width of the kernel  $\delta_i$ , and the weight of the OL  $w_i$ . Whereas  $w_i$  and  $\delta_i$  are determined by gradient descent learning [42] and fuzzy C-means (FCM) technique. Equations (22) and (23) represent the centre of the RBF ( $c_{ik}$ ) and the kernel width ( $\delta_i$ ), respectively.

$$c_{ik} = \frac{\sum_{j=1}^n \mu_{ji} x_{jk}}{\sum_{j=1}^n \mu_{ji}} \tag{22}$$

$$\delta_i = \frac{\sum_{j=1}^n \mu_{ji} \|x_j - c_i\|^2}{\sum_{j=1}^n \mu_{ji}} \tag{23}$$

$$Let \tilde{x}_i = \varphi(\|x - c_i\|), i = 1, 2, \dots, M \tag{24}$$

$$\tilde{x} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_M]^T \tag{25}$$

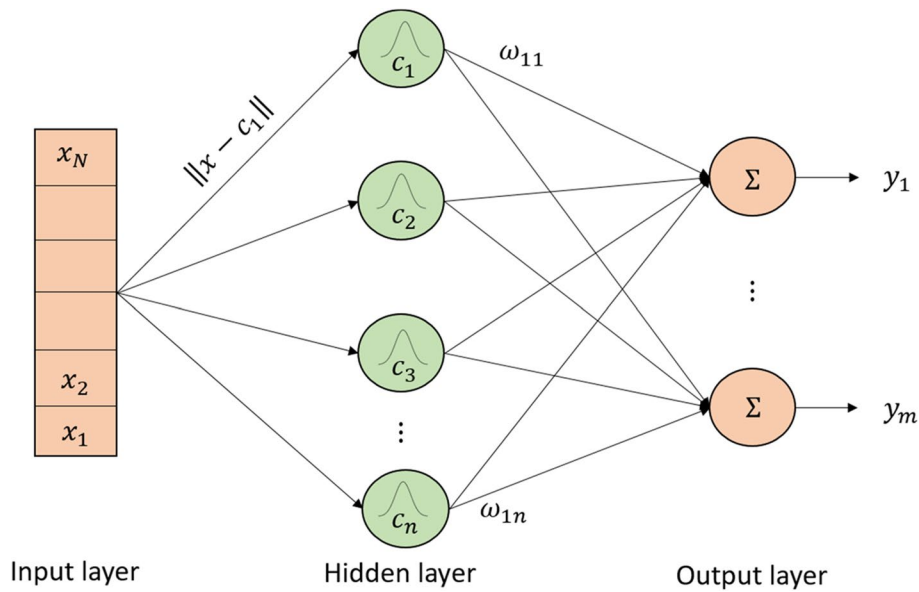
where, n represents the training US image samples,  $M$  represents the Total HL nodes, and  $\mu_{ji}$  represents the outcome of FCM. The input sample is mapped to the new space  $f : R^d \rightarrow R^M$ , and the transformation from the IL to the HL creates a nonlinear mapping. The network operation is expressed in Eq. (26) and let’s assume  $p = [w_1, w_2, \dots, w_M]^T$ .

$$y = p^T \tilde{x} \tag{26}$$

The network’s output can be transformed into a linear model by calculating the RBF’s HL, as illustrated in Eq. (26).

**Fast-RBFNN**

The RBF linear model’s  $\epsilon$ -insensitive loss function (ILF) is introduced through the fast-RBF search principle. The  $\epsilon$ -ILF refers to a loss function that reduces sensitivity to noise or variations in data, enabling more important features to be identified. Unlike common loss functions such



**Fig. 5** RBFNN Architecture

as cross-entropy, which aim to minimize overall prediction error, this function prioritizes robustness, making it more suitable for medical imaging tasks. By determining the significance of the constraint term  $\epsilon$ , the optimization seeks to minimize the  $\epsilon$ -ILF. The RBFNN optimization model is constructed using the Gaussian kernel and incorporates the structural risk term with large-sample processing. The structural risk term acts as a regularization component in the loss function, controlling model complexity. It helps mitigate overfitting by penalizing excessively complex models, ensuring better generalization to new data, and enhancing overall performance. The processes are detailed below.

1. Equations (22) and (23) yield the values of  $c_i$  and  $\delta_i$ , respectively, while Eq. (5) yields the model input  $\tilde{x}$ .
2. The loss function unaffected by  $\epsilon$  is presented. The definition of the  $\epsilon$ -ILF is as follows:

$$L^\epsilon(x, y, f) = |y - f(x)|_\epsilon = \max(0, |y - f(x)| - \epsilon), \quad x \in R^d, y \in R \quad (27)$$

Equation (26)'s linear model can have its equivalent  $\epsilon$ -ILF written as

$$\sum_{i=1}^n |y_i^o - y_i|_\epsilon = \sum_{i=1}^n \max(0, |y_i^o - y_i| - \epsilon) = \sum_{i=1}^n \max(0, |p^T \tilde{x}_i - y_i| - \epsilon) \quad (28)$$

The output value of the neural network is denoted by  $y_i^o$ , while  $y_i$  represents the actual output value. When the constraints of  $p^T \tilde{x}_i - y_i < \epsilon$  and  $y_i - p^T \tilde{x}_i < \epsilon$  are not always satisfied, the relaxation factors  $\xi_i$  and  $\xi_i^*$  are added.

$$\begin{cases} y_i - p^T \tilde{x}_i < \epsilon + \xi_i, \xi_i \geq 0 \\ p^T \tilde{x}_i - y_i < \epsilon + \xi_i^*, \xi_i^* \geq 0 \end{cases} \quad (29)$$

To minimize the value of the  $\epsilon$ -ILF, represented by Eq. (8), this algorithm exists. The  $\epsilon$ -insensitive parameter's value has a direct impact on the precision of the models. Therefore, the optimization problem incorporates the parameter  $\lambda$  and utilizes  $\epsilon$  as the constraint term. When the optimization problem is coupled with Eq. (29), the resulting expression remains the same and it is represented in Eq. (27)

$$\min 2\lambda\epsilon + \frac{\lambda}{\mu n} \sum_{i=1}^n (\xi_i^2, \xi_i^{*2}), \text{ s.t. } \begin{cases} y_i - p^T \tilde{x}_i < \epsilon + \xi_i \\ p^T \tilde{x}_i - y_i < \epsilon + \xi_i^* \end{cases} \quad (30)$$

The condition that  $\xi_i, \xi_i^* \geq 0$  is automatically met, and  $\mu$  is the equilibrium factor.

3. Introducing kernel functions and structural risk items: A support vector machine (SVM) is one method for minimizing structural risk [43, 44]. This work presents a strategy for learning how to design a SVM and incorporating a regularization to reduce algorithmic risk. SVMs use the kernel strategy to enhance the computational capabilities of linear learners. The proposed technique in this paper also involves the use of a kernel function [45]. Employing the regularization and kernel function can help to represent the optimization problem and it is given in Eq. (28).

$$\min_{p, \varepsilon, \xi_i, \xi_i^*} \|p\|^2 + 2\lambda\varepsilon + \frac{\lambda}{\mu n} \sum_{i=1}^n (\xi_i^2 + \xi_i^{*2}), s.t. \begin{cases} y_i - p^T \varphi(\tilde{x}_i) < \varepsilon + \xi_i, \\ p^T \varphi(\tilde{x}_i) - y_i < \varepsilon + \xi_i^*, \\ i = 1, 2, \dots, n \end{cases} \quad (31)$$

4. Formula derivation: The Lagrangian function can be written as a function of time when the Lagrange multiplier is introduced.

$$L = \|p\|^2 + 2\lambda\varepsilon + \frac{\lambda}{\mu n} \sum_{i=1}^n (\xi_i^2 + \xi_i^{*2}) + \sum_{i=1}^n \alpha_i (y_i - p^T \varphi(\tilde{x}_i) - \varepsilon - \xi_i) + \sum_{i=1}^n \alpha_i^* (p^T \varphi(\tilde{x}_i) - y_i - \varepsilon - \xi_i^*) \quad (32)$$

The related dual problem of Eq. (29) has the following matrix form:

$$\left\{ \max [\alpha^T \alpha^{*T}] \begin{bmatrix} \frac{2}{\lambda} y \\ -\frac{2}{\lambda} y \end{bmatrix} - [\alpha^T \alpha^{*T}] \tilde{K} \begin{bmatrix} \alpha \\ \alpha^* \end{bmatrix}, s.t. [\alpha^T \alpha^{*T}] \mathbf{1} = 1, \alpha, \alpha^* \geq 0 \right. \quad (33)$$

where  $\alpha, \alpha^*$  and  $\tilde{K}$  represents Lagrange coefficients and kernel function.

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}, \alpha^* = \begin{bmatrix} \alpha_1^* \\ \vdots \\ \alpha_n^* \end{bmatrix} \quad (34)$$

$$\tilde{K} = [\tilde{k}(\tilde{x}_i, \tilde{x}_j)] = \begin{bmatrix} K + \frac{\mu n}{\lambda} I & -K \\ -K & K + \frac{\mu n}{\lambda} I \end{bmatrix} \quad (35)$$

where  $K$  represents the Gaussian kernel. The solution's variable values are

$$\begin{cases} p = \lambda \sum_{i=1}^n (\alpha_i - \alpha_i^*) \varphi(\tilde{x}_i) \\ \xi_i = \alpha_i \mu n \\ \xi_i^* = \alpha_i^* \mu n. \end{cases} \quad (36)$$

5. Classification: The classifier function is illustrated in the Eq. (35)

$$y = \text{sign}(p^T \varphi(\tilde{x}_{\text{test}})) = \lambda \sum_{i=1}^n (\alpha_i - \alpha_i^*) \varphi^T(\tilde{x}_i) \varphi(\tilde{x}_{\text{test}}) = \lambda \sum_{i=1}^n (\alpha_i - \alpha_i^*) \tilde{K}(\tilde{x}_i, \tilde{x}_{\text{test}}) \quad (37)$$

A fetal brain, abdomen, femur, thorax, maternal cervix and other is associated with a y-value that falls anywhere from 0 to 5.

## Result and discussion

### Experimental setup

A novel methodology for identifying fetal organs in US images is introduced, utilizing the Zendo dataset. After pre-processing, features are extracted using GLCM techniques. These features are then inputted into conventional DL models, including ANN, CNN, RBF) and our proposed fast RBFNN. The experimental environment configuration setting used in this research is displayed in Table 4.

### Evaluation metrics

Assessing DL model performance is critical, achieved through various metrics such as Accuracy, Specificity, Sensitivity, Precision, F1 score, and Matthews Correlation Coefficient (MCC), False Rejection Rate (FRR) and False Acceptance Rate (FAR). Table 5 details the formula for each metric. In these formulas,  $TP$  represents True

**Table 4** Experimental environment

Model	Accuracy
Operating System	Windows 11
CPU	Intel(R) Xeon(R) Gold 6330
GPU	NVIDIA GeForce RTX
RAM	32 GB
Programming Language	Python
Environment	Anaconda – JupyterNotebeook
Proposed	96.37

**Table 5** Evaluation metrics

Metrics	Formula
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Specificity	$\frac{TN}{TN+FP}$
Sensitivity	$\frac{TP}{TP+FN}$
Precision	$\frac{TP}{TP+FP}$
F1 score	$\frac{2TP}{2TP+FP+FN}$
MCC	$\frac{TP-FN}{TP+FN}$
FRR	$\frac{FN}{TP+FN}$
FAR	$\frac{FP}{TN+FP}$

Positives, *TN* represents True Negatives, *FP* represents False Positives, and *FN* represents False Negatives in the classification of fetal organs in US images.

**Experimental outcome**

Table 6 gives the performance metrics attained by the convention and proposed DL models using GLCM features. Our proposed model exhibits the highest performance metrics, with Accuracy, Specificity, Sensitivity, Precision, F1 score, and MCC at 96.37%, 97.08%, 95.67%, 97.02%, 96.34%, and 92.75%, respectively, and minimal FRR and FAR of 4.33 and 2.92 using GLCM features.

To enhance model performance, essential features are selected using optimized algorithms like PSO, GWO, and a hybrid PSO-GWO approach. The performance of both conventional and proposed DL models, utilizing features selected by these algorithms, is presented in Tables 7, 8 and 9. The combination of hybrid PSOGWO-selected features with our proposed model yields the most satisfactory performance, achieving 98.07% Accuracy, 98.12% Specificity, 98.03% Sensitivity, 98.18% Precision, 98.10% F1 score, and 96.15% MCC.

Figures 6 and 7 illustrate a comparison of DL models based on positive and negative metrics using various feature extraction techniques, including GLCM, PSO, GWO, and hybrid approaches. The figures highlight that the proposed model, which utilizes hybrid features, outperforms other DL models and feature extraction techniques in terms of both positive and negative metrics. This demonstrates the superior performance and effectiveness of the proposed framework.

The speckle noise removal is an important pre-processing step. The outcome of the DL models using Hybrid PSOGWO-selected features gives better results in fetal

organ classification. To demonstrate the importance of the pre-processing step, the DL model is evaluated using the Hybrid PSOGWO-selected features with the input image containing speckle noise. The outcome of the DL model using Hybrid PSOGWO-selected features without noise removal is presented in Table 10. By comparing Tables 9 and 10, the DL model achieves the highest positive metrics and lower negative metrics on input images without noise.

The performance of the proposed model in classifying each organ from US images is evaluated, with results summarized in Table 11. The proposed model exhibits high accuracy across various organs: 97.69% for the brain, 98.15% for the abdomen, 99.07% for the femur, 97.69% for the thorax, 97.22% for maternal cervix, and 98.61% for other organs. In addition to accuracy, other metrics for each organ are also provided in Table 11, offering a comprehensive assessment of the proposed model's performance across different organs in fetal imaging.

The time efficiency of each stage of the proposed methodology on test data is evaluated and presented in Table 12. Pre-processing and feature extraction using GLCM require 4 s and 11 s, respectively. Subsequently, feature selection using the hybrid PSO-GWO algorithm consumes a significant time of 5 min and 29 s. For classification, ANN takes 38 s, while our proposed model takes 57 s, although it is faster than CNN and RBF models. This assessment provides insight into the computational time required by the proposed method. Overall, the finalized methodology, comprising pre-processing, GLCM feature extraction, PSOGWO-based feature selection, and fast RBF classification, takes 6 min and 41 s to process the test data.

In assessing the effectiveness of the proposed model, a comparative analysis with existing models from recent

**Table 6** DL model performance on fetal US image classification using GLCM features

Model	Accuracy	Specificity	Sensitivity	Precision	F1	MCC	FRR	FAR
ANN	92.59	93.88	91.35	93.92	92.62	85.22	8.65	6.12
CNN	93.06	94.25	91.87	94.18	93.01	86.13	8.13	5.75
RBF	94.98	95.25	94.71	95.16	94.93	89.96	5.29	4.75
Proposed	96.37	97.08	95.67	97.02	96.34	92.75	4.33	2.92

**Table 7** DL model performance on fetal US image classification using PSO selected features

Model	Accuracy	Specificity	Sensitivity	Precision	F1	MCC	FRR	FAR
ANN	94.06	94.70	93.44	94.74	94.08	88.12	6.56	5.30
CNN	96.06	97.06	95.08	97.02	96.04	92.14	4.92	2.94
RBF	96.84	97.68	96.00	97.65	96.82	93.68	4.00	2.32
Proposed	97.54	98.31	97.37	98.28	97.83	95.68	2.63	1.69

**Table 8** DL model performance on fetal US image classification using GWO selected features

Model	Accuracy	Specificity	Sensitivity	Precision	F1	MCC	FRR	FAR
ANN	94.91	95.52	94.29	95.47	94.88	89.82	5.71	4.48
CNN	95.22	96.39	94.08	96.42	95.24	90.46	5.92	3.81
RBF	96.76	97.23	96.29	97.19	96.74	93.52	3.71	2.77
Proposed	97.84	98.16	97.52	98.13	97.82	95.68	2.48	1.84

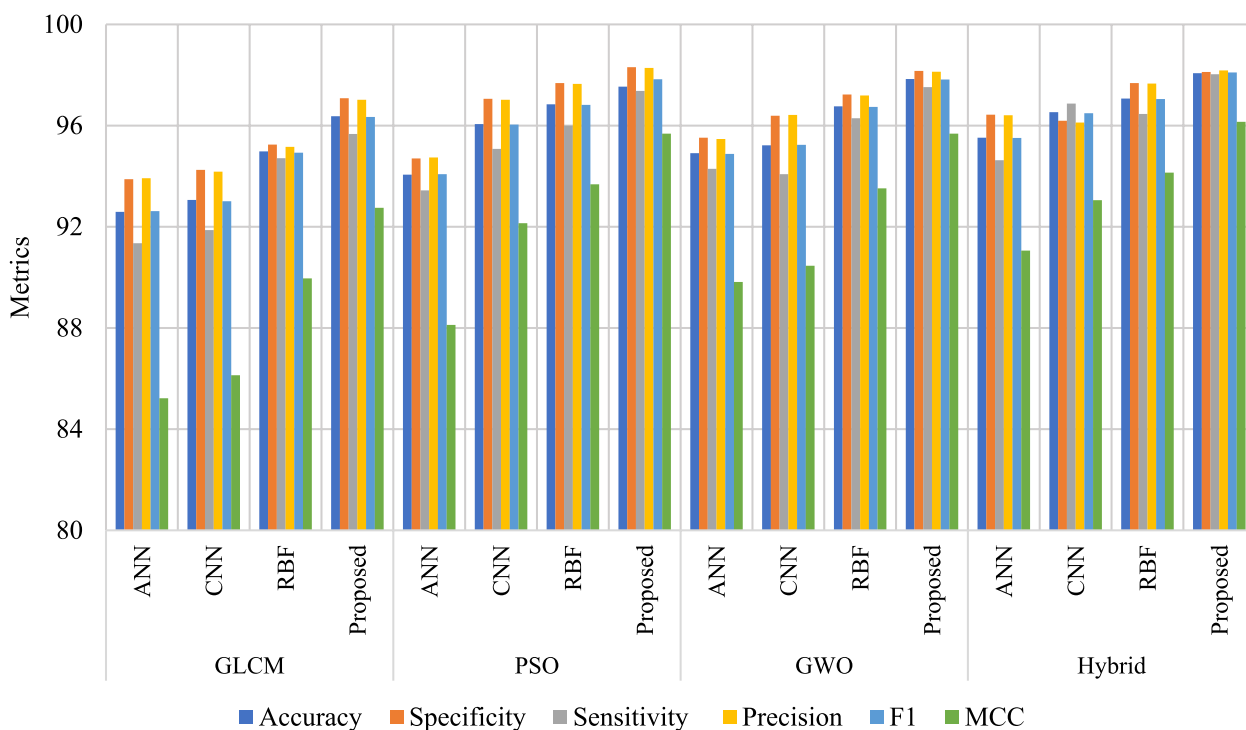
**Table 9** DL model performance on fetal US image classification using Hybrid PSOGWO selected features

Model	Accuracy	Specificity	Sensitivity	Precision	F1	MCC	FRR	FAR
ANN	95.52	96.43	94.63	96.41	95.51	91.06	5.37	3.57
CNN	96.53	96.19	96.87	96.12	96.49	93.05	3.13	2.05
RBF	97.07	97.68	96.46	97.66	97.05	94.14	3.54	2.32
Proposed	98.07	98.12	98.03	98.18	98.10	96.15	1.97	1.88

research was conducted, as presented in Table 13. Analysis revealed that several references, namely [46, 48], and [49], evaluated their models solely based on accuracy, without considering other essential metrics. In contrast, our proposed model offers comprehensive metrics. The highest accuracy reported in previous research is 96.85% and 96% from studies [47, 48]. The proposed model

achieves 98.07%, surpassing the existing best accuracy. For sensitivity and precision, the highest scores reported are from references [24, 47], with sensitivity values of 96.28% and 96.66%, and precision values of 94.02% and 97.12%. The proposed model yields 98.12% sensitivity and 98.18% precision, which is significantly better than the existing results. The MCC values reported

### Positive Metrics Comparison of DL models using Various Features



**Fig. 6** Evaluation of the DL model's positive metrics by employing various features

### Negative Metrics Comparison of DL models using Various Features

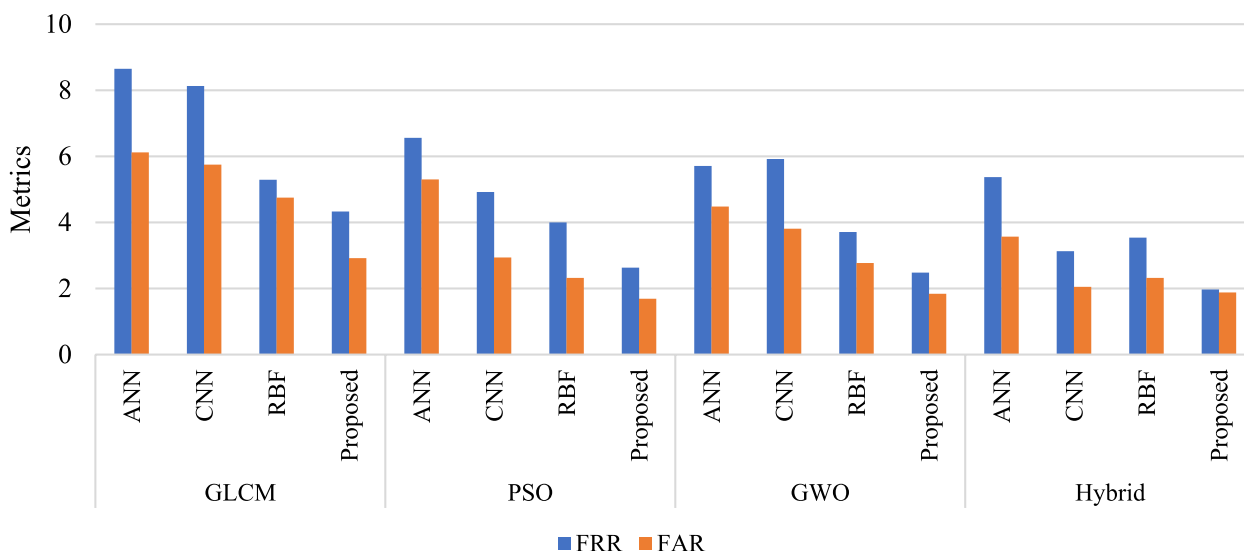


Fig. 7 Evaluation of the DL model's negative metrics by employing various features

Table 10 DL model performance on fetal US image classification using Hybrid PSOGWO selected features without noise removal

Model	Accuracy	Specificity	Sensitivity	Precision	F1	MCC	FRR	FAR
ANN	94.47	95.24	93.25	95.47	94.35	90.01	6.24	4.83
CNN	95.23	95.03	95.65	94.79	95.22	92.55	3.98	3.62
RBF	95.89	96.14	95.98	96.24	96.11	93.69	4.57	3.46
Proposed	96.54	97.2	96.89	96.88	96.88	95.78	2.68	2.49

Table 11 Evaluation of the proposed methodology for each image classification

	Accuracy	Specificity	Sensitivity	Precision	F1	MCC	FRR	FAR
Brain	97.69	98.10	97.30	98.18	97.74	95.37	2.70	1.90
Abdomen	98.15	99.04	97.32	99.09	98.20	96.31	2.68	0.96
Femur	99.07	99.06	99.09	99.09	99.09	98.14	0.91	0.94
Thorax	97.69	97.20	98.17	97.27	97.72	95.37	1.83	2.80
Maternal Cervix	97.22	96.30	98.15	96.36	97.25	94.46	1.85	3.70
Other	98.61	99.06	98.18	99.08	98.63	97.22	1.82	0.94
Average	98.07	98.12	98.03	98.18	98.10	96.15	1.97	1.88

in research [19, 24] are 74.8% and 94.19%, whereas the proposed model achieves an MCC of 96.15%. Additionally, for the F1 score, the proposed model outperforms existing values, which range from 79 to 97%. Overall,

the comparative analysis highlights the effectiveness and superiority of our proposed model over existing approaches, showcasing its potential for advancing fetal organ identification in US imaging research.



**Table 12** Analysis of operational efficiency on test data

Steps	Method	Time-Consuming
Pre-processing	-	4 s
Feature Extraction	GLCM	11 s
Feature Selection	PSO	2 min 20 s
	GWO	3 min 12 s
	PSOGWO	5 min 29 s
Classifiers	ANN	38 s
	CNN	1 min 5 s
	RBF	1 min 12 s
	Proposed	57 s

**Table 13** Comparison of the proposed model with existing works

Ref	Accuracy	Sensitivity	Precision	MCC	F1 score
[19]	79.4	79.45	78.95	74.8	79.1
[46]	93.6	-	-	-	-
[24]	95.69	96.28	94.02	94.19	95.08
[47]	96.85	96.66	97.12	-	96.88
[48]	96	-	-	-	-
[49]	94	-	-	-	-
Our	98.07	98.12	98.18	96.15	98.10

**Discussion**

The proposed model, using hybrid PSO-GWO features, delivers excellent results and outperforms existing models such as ANN, CNN, and RBF. The model achieves accuracy, sensitivity, precision, MCC, and F1 scores of 98.07%, 98.12%, 98.18%, 96.15%, and 98.10%, respectively. Table 13 demonstrates that the proposed model also performs better than existing research. The  $\epsilon$ -insensitive loss function ( $\epsilon$ -ILF) and advanced feature selection contribute to the model’s improved performance by handling noisy images more effectively. Additionally, structural risk minimization helps reduce overfitting. Another reason for the improved results is the selection of more suitable features using the hybrid PSO-GWO approach. Table 11 presents the model’s performance in classifying each organ, indicating that the model accurately identifies all organs. The outcomes of the proposed framework suggest that the model is suitable for real-time deployment and provides promising results in identifying fetal organs from US images.

However, the proposed framework also presents challenges. The execution time for feature extraction, feature selection using the hybrid model, and classification is 11 s, 5 min 29 s, and 57 s, respectively. The complexity of the framework increases the execution time. Analysis

of Tables 9 and 10, which compare the hybrid DL model with and without noisy images, reveals that the model performs better without noisy images. This indicates that the proposed model still relies heavily on pre-processing steps to achieve optimal performance.

**Conclusion**

The research successfully designs a novel methodology using hybrid optimized feature selection and the Fast RBFNN model for maternal fetal US plane classification. US images are the safest technique for fetal and maternal health monitoring. These images are processed and features are extracted using the GLCM technique. The most important features are selected by the PSOGWO algorithm, as well as PSO and GWO. The selected features from each algorithm are then fed into DL classifiers such as ANN, CNN, RBF, and Fast RBFNN. All combinations of output features from GLCM, PSO, GWO, and PSOGWO with DL models are tested. Experimental results show that the proposed model performs better on hybrid PSOGWO features. The accuracy attained by the proposed model is 98.07%, 97.84%, 97.54%, and 96.37% on PSOGWO, GWO, PSO, and GLCM features, respectively. Next, the model efficiency is compared using time consumption. The proposed DL model takes only 57 s to classify the test data, indicating its suitability for real-time applications. A limitation of the research is the use of data from the same dataset for both training and testing. To better determine the model’s effectiveness, it is necessary to compare it with additional datasets. Future work will involve evaluating the model with diverse datasets and exploring its deployment for clinical use in real-time applications.

**Authors’ contributions**

S. Rathika- Paper writeup, Literature review, Comparative study K. Mahendran- Methodology, Outline, Software tool H. Sudarsan- Paper writeup, Literature review S. Vijay Anand- Manuscript preparation, Language correction.

**Funding**

The authors have no financial support or no funding from anybody on this article.

**Data availability**

No datasets were generated or analysed during the current study.

**Declarations**

**Ethics approval and consent to participate**

This research not performed with any Human or Animal and not affected environment directly or indirectly. All authors are included in the manuscript are participated, and there is no other participation.

**Consent for publication**

Hereby, we confirm that, the article titled as “Novel Neural Network Classification of Maternal Fetal Ultrasound Planes through Optimized Feature Selection” has not been published previously and it is not under consideration elsewhere.

**Competing interests**

The authors declare no competing interests.

Received: 13 May 2024 Accepted: 4 October 2024

Published online: 18 December 2024

**References**

- Barnett SB, Maulik D. Guidelines and recommendations for safe use of Doppler ultrasound in perinatal applications. *J Maternal-Fetal Med.* 2001;10(2):75–84.
- Whitworth M, Bricker L, Mullan C. Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst Reviews* 2015;7:1–78.
- Queyam A, Bin SK, Pahuja, Singh D. Non-invasive feto-maternal well-being monitoring: a review of methods. *J Eng Sci Technol Rev.* 2017;6(5):7–14.
- Lockwood CJ, Magriples U. Prenatal care: Second and third trimesters. Published online in uptodate.com. 2022.
- Araujo E Júnior, Santana EFM, Martins WP, Elito J Júnior, Ruano R, Pires CR, Filho SMZ. Reference charts of fetal biometric parameters in 31,476 Brazilian singleton pregnancies. *J Ultrasound Med.* 2014;33(7):1185–91.
- Nelson TR, Pretorius DH. Three-dimensional ultrasound of fetal surface features. *Ultrasound Obstet Gynecol.* 1992;2(3):166–74.
- Zhang L, Chen S, Chin CT, Wang T, Li S. Intelligent scanning: Automated standard plane selection and biometric measurement of early gestational sac in routine ultrasound examination. *Med Phys.* 2012;39(8):5015–27.
- Shahrivar RY, Karami F, Karami E. Enhancing fetal anomaly detection in ultrasonography images: a review of machine learning-based approaches. *Biomimetics.* 2023;8(7):519.
- Stanton K, Mwanri L. Global maternal and child health outcomes: the role of obstetric ultrasound in low resource settings. *J Prev Med.* 2013;1(3):22–9.
- Jena B, Nayak GK, Saxena S. Convolutional neural network and its pre-trained models for image classification and object detection: a survey. *Concurr Comput: Pract Exp.* 2022;34(6):e6767.
- Saba L, Biswas M, Kuppili V, Godia EC, Suri HS, Edla DR, Omerzu T, et al. The present and future of deep learning in radiology. *Eur J Radiol.* 2019;114:14–24.
- Xie Q, Faust K, Van Ommeren R, Sheikh A, Ugljesa Djuric, and Phedias Diamandis. Deep learning for image analysis: personalizing medicine closer to the point of care. *Crit Rev Clin Lab Sci.* 2019;56(1):61–73.
- Sridar P, Kumar A, Quinton A, Nanan R, Kim J. Decision fusion-based fetal ultrasound image plane classification using convolutional neural networks. *Ultrasound Med Biol.* 2019;45(5):1259–73.
- Szegedy C, et al. Going deeper with convolutions, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston. 2015. pp. 1–9. <https://doi.org/10.1109/CVPR.2015.7298594>.
- He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: 2016. pp. 770–8. <https://doi.org/10.1109/CVPR.2016.90>.
- Yu Z, Tan E-L, Ni D, Qin J, Chen S, Li S, Lei B, Wang T. A deep convolutional neural network-based framework for automatic fetal facial standard plane recognition. *IEEE J Biomed Health Inf.* 2017;22(3):874–85.
- Zhao Z, Alzubaidi L, Zhang J, Duan Y, Gu Y. A comparison review of transfer learning and self-supervised learning: definitions, applications, advantages and limitations. *Expert Syst Appl.* 2023;122807:1–38.
- Chen H, Ni D, Qin J, Li S, Yang X, Wang T, Pheng Ann Heng. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. *IEEE J Biomedical Health Inf.* 2015;19(5):1627–36.
- Rauf F, Khan MA, Bashir AK, Jabeen K, Hamza A, Alzahrani AI, Alalwan N, Masood A. Automated deep bottleneck residual 82-layered architecture with bayesian optimization for the classification of brain and common maternal fetal ultrasound planes. *Front Med.* 2023;10:1330218.
- Lasala A, Fiorentino MC, Bandini A, Moccia S. FetalBrainAwareNet: bridging GANs with anatomical insight for fetal ultrasound brain plane synthesis. *Comput Med Imaging Graph.* 2024;116:102405.
- Montero A, Bonet-Carne E, Burgos-Artizzu XP. Generative adversarial networks to improve fetal brain fine-grained plane classification. *Sensors.* 2021;21(23):7975.
- Ashkani A, Vahid MG, Oghli A, Shabanzadeh N, Sirjani A, Akhavan I, Arabi SH, Taheri MS, Tarzamni MK. Morteza Sane'i Taheri, and Mohammad Kazem Tarzamni. Fast and accurate U-net model for fetal ultrasound image segmentation. *Ultrason Imaging.* 2022;44(1):25–38.
- Gao Z, Tan G, Wang C, Lin J, Pu B, Li S, Li K. Graph-enhanced ensembles of multi-scale structure perception deep architecture for fetal ultrasound plane recognition. *Eng Appl Artif Intell.* 2024;136:108885.
- Krishna T, Bala, Kokil P. Standard fetal ultrasound plane classification based on stacked ensemble of deep learning models. *Expert Syst Appl.* 2024;238:122153.
- Migliorelli G, Fiorentino MC, Cosmo MD. Francesca Pia Villani, Adriano Mancini, and Sara Moccia. On the use of contrastive learning for standard-plane classification in fetal ultrasound imaging. *Comput Biol Med.* 2024;174:108430.
- Krishna TB, Kokil P. Automated classification of common maternal fetal ultrasound planes using multi-layer perceptron with deep feature integration. *Biomed Signal Process Control.* 2023;86:105283.
- Available Online. <https://zenodo.org/records/3904280&#8216>.
- Luke J, Rajkumar J, Balaji M. Impact of image size on accuracy and generalization of convolutional neural networks. *Int J Res Anal Rev.* 2019;6(1):70–80.
- Yan J, Lin S, Kang SB, Tang X. Learning the Change for Automatic Image Cropping, 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland: 2013. pp. 971–78. <https://doi.org/10.1109/CVPR.2013.130>.
- Ali PJ, Faraj RH, Koya E, Ali PJ, Faraj RH. Data normalization and standardization: a technical report. *Mach Learn Tech Rep.* 2014;1(1):1–6.
- Abdani S, Raihanah MA, Zulkifley, Nuraisyah Hani Z. Undersampling and oversampling strategies for convolutional neural networks classifier, Proceedings of the 6th International Conference on Electrical, Control and Computer Engineering. 2022. pp. 1129–37.
- Rebuffi S-A, Gowal S, Calian DA, Stimberg F, Wiles O, Timothy A. Mann. Data augmentation can improve robustness. *Adv Neural Inf Process Syst.* 2021;34:29935–48.
- Mohanaiah P, Sathyanarayana P, GuruKumar L. Image texture feature extraction using GLCM approach. *Int J Sci Res Publications.* 2013;3(5):1–5.
- Muntasa A. Multi Distance and Angle models of the Gray Level Co-occurrence Matrix (GLCM) to extract the Acute Lymphoblastic Leukemia (ALL) images. *Int J Intell Eng Syst.* 2021;14:6.
- Haralick RM. Karthikeyan Shanmugam, and its' Hak Dinstein. Textural features for image classification. *IEEE Trans Syst Man Cybernet.* 1973;6:610–21.
- Shen C, Zhang K, Tang J. A covid-19 detection algorithm using deep features and discrete social learning particle swarm optimization for edge computing devices. *ACM Trans Internet Technol (TOIT).* 2021;22(3):1–17.
- Jain M, Saihjpal V, Singh N, Bir Singh S. An overview of variants and advancements of PSO algorithm. *Appl Sci.* 2022;12(17):8392.
- Tu C-J, Chuang L-Y, Jun-Yang C, Cheng-Hong Y. Feature selection using PSO-SVM. *Int J Comput Sci.* 2007;33:1.
- Mirjalili S, Lewis A. Grey wolf optimizer. *Adv Eng Softw.* 2014;69:46–61.
- Hatta NM, Zain AM, Sallehuddin R, Shayfull Z, Yusoff Y. Recent studies on optimisation method of Grey Wolf Optimiser (GWO): a review (2014–2017). *Artificial Intell Rev.* 2019;52:2651–83.
- Han H-G, Jun-Fei Q. Adaptive computation algorithm for RBF neural network. *IEEE Trans Neural Networks Learn Syst.* 2011;23(2):342–7.
- Haji SH, Adnan Mohsin Abdulazeez. Comparison of optimization techniques based on gradient descent algorithm: a review. *PalArch's J Archaeol Egypt/Egyptology.* 2021;18(4):2715–43.
- Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods. Cambridge University Press; 2000. <https://doi.org/10.1017/CBO9780511801389>.
- Chicco D. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genomics.* 2020;21:1–13.
- Vujović Ž. Classification model evaluation metrics. *Int J Adv Comput Sci Appl.* 2021;12(6):599–606.
- Burgos-Artizzu XP, Coronado-Gutiérrez D, Brenda Valenzuela-Alcaraz, Elisenda Bonet-Carne, Elisenda Eixarch, Fatima Crispi, and Eduard Gratacós. Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Sci Rep.* 2020;10(1):10200.

47. Thomas S, Harikumar S. An ensemble deep learning framework for foetal plane identification. *Int J Inform Technol.* 2014;16:1377–86.
48. Al-Razgan M, Ali YA. Enhancing Fetal Medical Image Analysis through attention-guided convolution: a comparative study with established models. *J Disabil Res.* 2024;3(2):20240005.
49. Srivastava S, Vidyarthi A, Jain S. A highly densed deep neural architecture for classification of the multi-organs in fetal ultrasound scans. *Neural Comput Appl* 2023:1–15.

### **Publisher's note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.