Neural Computing and Applications
https://doi.org/10.1007/s00521-023-09148-x

S.I. : INTELLIGENT SYSTEMS IN BIOMEDICAL AND HEALTHCARE INFORMATICS

A highly densed deep neural architecture for classification of the multi-organs in fetal ultrasound scans

Somya Srivastava[1] · Ankit Vidyarthi[1] · Shikha Jain[1]

Received: 10 April 2023 / Accepted: 17 October 2023

Abstract

Artificial intelligence (AI) makes a substantial contribution to decision-making in many intricate application areas of the medical sciences. One such application is *organ classification* in maternal–fetal ultrasound scans using sophisticated AI methodology like deep neural networks for better analysis. In this paper, we present a novel and highly dense deep neural architecture specifically designed for the multi-organ classification of fetal ultrasound scans. Our proposed approach introduces a unique combination of densely connected layers, convolution layers, and skip connections, tailored to accurately identify and classify multiple fetal organs simultaneously. To the best of our knowledge, this is the first study to address the comprehensive classification of multiple organs in fetal ultrasound images using such a highly dense neural network. The architecture is designed to capture both local and global features, critical for distinguishing intricate organ structures in ultrasound images. The model also minimizes the gradient loss for faster convergence of the model parameters in the training phase. Through extensive evaluation of a curated dataset of fetal ultrasound scans, we demonstrate that our novel architecture achieves superior classification accuracy—96.85%, precision—97.12%, recall—96.66%, F1-score—96.88%, and AUC-ROC score—97.27%, outperforming state-of-the-art methods in the context of multi-organ classification. Thus, the proposed highly dense deep neural architecture presents a promising avenue for enhancing fetal ultrasound imaging, bringing potential benefits to prenatal care, and contributing to improved neonatal outcomes.

Keywords Deep neural network · Medical image classification · Ultrasound imaging · Fetus analysis

1 Introduction

Ultrasound imaging is a common method used to visualize a developing fetus in the mother's womb. It is a valuable tool for monitoring fetal development and ensuring a healthy pregnancy. It is a non-invasive and safe procedure that uses high-frequency sound waves to create images of the fetus, which can be used to assess fetal growth and development, as well as identify any potential abnormalities. During an ultrasound for fetal imaging, a gel is applied to the mother's abdomen, and a transducer is placed on the gel to emit sound waves into the uterus. The sound waves bounce back off the developing fetus and are captured by the transducer, which converts the echoes into an image that can be viewed on a monitor.

Ultrasound imaging can provide valuable information about the fetal anatomy, including the size, shape, and position of the fetus, as well as the development of the organs, limbs, and other structures. It can also be used to monitor fetal movements and heart rate, as well as assess the amount of amniotic fluid surrounding the fetus. In addition to its diagnostic uses, ultrasound imaging can also be used to guide certain medical procedures, such as amniocentesis or fetal surgery.

There are different types of ultrasound imaging that may be used during pregnancy, including:

– Transabdominal ultrasound: This is the most common type of ultrasound used during pregnancy. A

✉ Ankit Vidyarthi
ankit.vidyarthi@jiit.ac.in

Somya Srivastava
18403019@mail.jiit.ac.in

Shikha Jain
shikha.jain@jiit.ac.in

[1] Department of Computer Science Engineering and Information Technology, Jaypee Institute of Information Technology, Noida, India

transducer, which is a handheld device, is moved over the mother's abdomen to create an image of the fetus.
– Transvaginal ultrasound: In some cases, a transvaginal ultrasound may be performed, particularly in early pregnancy. This involves inserting a small, wand-shaped transducer into the vagina to obtain images of the uterus and fetus.
– 3D and 4D ultrasound: These advanced imaging techniques can provide a more detailed view of the fetus and are often used for diagnostic purposes.

Ultrasound imaging during pregnancy can help doctors assess the health and development of the fetus, including checking for multiple fetuses, measuring fetal growth, evaluating the placenta and amniotic fluid levels, and identifying potential birth defects or abnormalities.

It is important to note that while ultrasound imaging is generally considered safe, it should only be performed by trained medical professionals and used judiciously to minimize any potential risks to the mother and fetus.

During the first trimester of pregnancy, which lasts from conception to the end of week 13, the fetal organs begin to develop. At this early stage, the fetus is still very small, and most of its organs are still in the early stages of formation. The overview of some of the basic organs that develop in the early stage is shown in Fig. 1 which is taken from one of the review articles published on USG [7]. Here are some of the key organs and structures that develop during the first trimester:

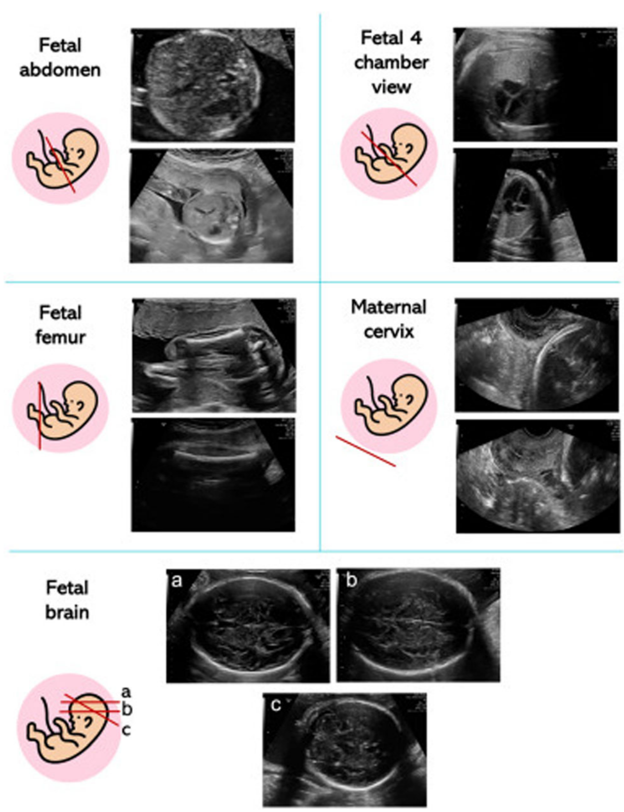

Fig. 1 Basic organs of the fetus in the early stage of formation [7]

1. Brain and nervous system: The brain and nervous system begins to form very early in the first trimester, with the neural tube forming around 3–4 weeks after conception.
2. Heart: The fetal heart begins to beat at around 6 weeks after conception, and the four chambers of the heart begin to develop soon after.
3. Lungs: The lungs begin to form in the first trimester, but they are not yet functional at this stage.
4. Liver: The liver begins to develop in the first trimester, and it is already functioning to some extent by the end of this stage.
5. Kidneys: The kidneys begin to form in the first trimester and begin to produce urine by around 10–12 weeks after conception.
6. Digestive system: The digestive system begins to form in the first trimester, with the development of the mouth, esophagus, stomach, and intestines.
7. Limbs: The fetal limbs begin to form in the first trimester, with the development of the arms and legs.

While the basic structures of many organs are present in the first trimester, the organs will continue to develop and mature throughout the rest of the pregnancy. The description of this is shown in Fig. 2.

Whereas, for early detection and automatic diagnosis of the fetus's health, AI tools and methodologies provide much valuable information. AI has shown great promise in the field of fetal organ analysis [1, 26]. AI algorithms can analyze large amounts of medical data, including ultrasound images, and provide an accurate and efficient analysis of fetal organs. One area where AI has been particularly useful is in the analysis of fetal brain development. AI algorithms can detect abnormalities in the brain structure and provide early detection of conditions such as hydrocephalus, neural tube defects, and ventriculomegaly. This can allow for early intervention and treatment, which can improve outcomes for babies.

AI is also being used to analyze fetal heart development. By analyzing ultrasound images of the fetal heart, AI algorithms can detect abnormalities such as heart defects and arrhythmias. This can help doctors make early diagnoses and develop treatment plans for the baby. But AI is also found useful in the analysis of fetal growth. AI algorithms can use ultrasound measurements to predict fetal weight and estimate the likelihood of complications such as macrosomia (large birth weight), which can increase the risk of difficult deliveries and neonatal complications.

Overall, AI has great potential to improve fetal organ analysis and provide better outcomes for babies. However,

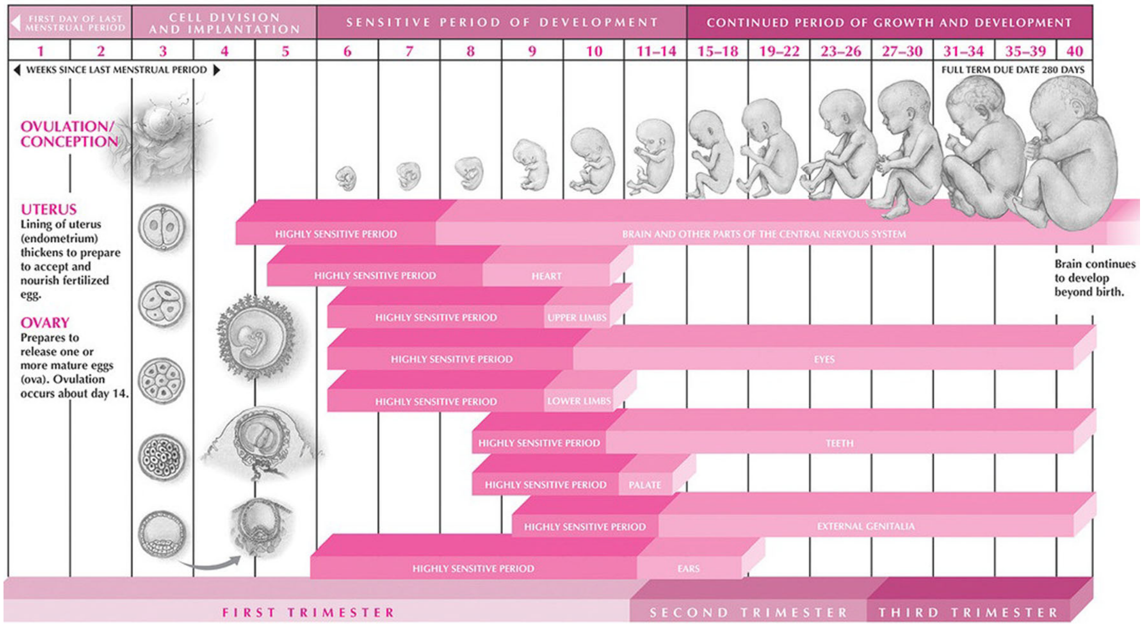

Fig. 2 Timeline for the fetus development

it is important to note that AI should always be used in conjunction with medical professionals, who can provide additional context and interpretation of the data.

The motivation behind using advanced AI methodologies and frameworks for the detection of fetal organs is to ensure the health and well-being of the mother and the developing fetus. There are several reasons why healthcare professionals seek to detect fetal organs, i.e., (i) to monitor fetal growth and development, ensuring that it is developing as it should, (ii) to detect potential health issues for early intervention and treatment, (iii) to prepare for safe delivery by knowing the location and health of fetal organs, (iv) to provide reassurance to the mother that the baby is developing normally and alleviate any anxieties or concerns she may have, and (v) to facilitate appropriate care to both the mother and the developing fetus.

In summary, detecting fetal organs is essential for ensuring the health and well-being of both the mother and the baby, allowing for early intervention and appropriate care.

To prioritize the fulfillment of the above motivations, this paper proposed the newly designed dense neural network that can be used for the classification of maternal–fetal organs in the first trimester. The main contributions of the work are summarized as follows:

1. Proposed a newly designed dense architecture that incorporates feature extraction blocks, which are designed to capture informative and discriminative features from the input data. These blocks allow the model to learn hierarchical representations of the input, enabling it to capture both low-level and high-level features that are essential for accurate classification.
2. The proposed architecture includes dense connectivity, where each layer is connected to all subsequent layers. This design choice encourages feature reuse across the network, enhancing the flow of information and improving the model's ability to retain critical information throughout the depth of the network.
3. To manage the computational complexity of the model, reduction blocks are introduced to downsample the feature maps and reduce their spatial dimensions.
4. The experimentations are performed on the highly competent Nvidia supercomputing system for model training and testing.

The main strength of the proposed highly densed deep neural architecture with feature extraction and reduction blocks lies in its ability to efficiently extract and retain informative features while effectively managing computational complexity. By incorporating dense connectivity and reduction blocks, the architecture achieves better feature reuse, improved generalization, and scalability, making it a robust and versatile solution for fetal organs classification in ultrasound images.

The rest of the paper is arranged in the subsequent section highlighting important modules as the background related work on the fetal organs is presented in Sect. 2. The description of the proposed architecture design with the important layer descriptions is presented in Sect. 3. In Sect. 4, the experimentation results of the proposed model using ultrasound scans are presented. Later, the comparative study is presented in Sect. 5. Finally, the paper gives the key summarization of the work and is presented in Sect. 6.

2 Related literature review

Fetal organ analysis is a crucial aspect of prenatal care, and various medical imaging techniques are used to analyze the development and health of the fetus's organs. Ultrasound is the most commonly used imaging technique for fetal organ analysis, and it can provide information about the size, shape, and function of organs such as the brain, heart, lungs, and kidneys. The use of deep learning for the segmentation and analysis of fetal organs has emerged as a promising area of research in recent years. Accurate identification and quantification of fetal organs can aid in the early detection of developmental abnormalities and improve prenatal care. This literature review explores the current state of research on deep learning algorithms and methodologies for fetal organ segmentation, classification, and analysis using ultrasound scans.

Deep learning has shown great potential for various medical imaging applications, including fetal organ classification, and can help diagnose fetal anomalies and malformations. The use of deep learning models in fetal organ classification involves training a neural network on a large dataset of labeled fetal ultrasound images. The neural network is then able to learn patterns and features from the images to accurately classify fetal organs. The majority of the studies focused on the fetal head, fetal brain, and heart segmentation, while a few also addressed the segmentation of other organs such as lungs, liver, and kidneys.

One common approach for fetal organ classification is to use a convolutional neural network (CNN), which is a type of neural network that is designed to process image data. CNNs consist of multiple layers of interconnected neurons that are able to automatically learn and extract features from images. In one of the recent articles, the authors explored the use of various CNN models to classify the common maternal–fetal ultrasound planes [1].

In recent years, besides ultrasound images, researchers have also explored the use of magnetic resonance imaging (MRI) for fetal organ analysis. MRI can provide detailed 3D images of fetal organs and structures, which can be useful for detecting abnormalities and monitoring fetal development. There has also been research on developing machine learning and computer vision algorithms for analyzing fetal organ images.

Another approach for fetal organ classification is to use a multi-task learning framework, where the neural network is trained to simultaneously classify multiple fetal organs. This approach can improve the accuracy of the classification task by leveraging shared features between different organs.

There has been a significant amount of research on fetal organ classification using deep learning in recent years [16]. This paper introduces the benchmark works and algorithms using the deep learning models for the automatic detection of the fetal head and the segmentation of standard planes in ultrasound images.

In the literature, there are several works that focus on the segmentation of the fetal organs and making proper assessments of the biometry parameters. One such work is proposed in [19] that presents a novel deep neural network for segmenting fetal organs in ultrasound images. The proposed method achieved high accuracy in detecting multiple organs. On a similar aspect, the work introduced in [6] used a multi-scale convolutional neural network for accurate segmentation of fetal abdominal organs in ultrasound images.

2.1 Literatures on fetal brain

Several studies have explored the use of deep learning for fetal brain segmentation from ultrasound scans. One such study was proposed in [11] where the authors introduced the method for automatic segmentation of fetal brain structures using a 2D convolutional neural network (CNN). The network was trained on a dataset of 10 fetal ultrasound volumes and achieved a mean Dice coefficient of 0.86 in segmenting the fetal brain structures.

Similarly, another study by the authors in [15, 31] proposed a deep learning-based method for fetal brain segmentation by an automatic segmentation tool for six-month child brain scans. Also, the work is reported on the quality assessment of the ultrasound images in [15].

In the literature, there were several research articles that used deep learning for fetal brain classification from ultrasound images. One study by the authors of [24] proposed a method for the automatic classification of fetal brain abnormalities using a convolutional neural network (CNN). The authors used the ultrasound video for the experimentation which resulted in an accuracy of 96%.

Similarly, another study proposed a deep learning architecture for the recognition of fetal brain standard scan planes in 2D ultrasound images [22]. The authors have used the two dataset studies having 455 scans in total for validation of the work. Using deep architecture, the authors claimed to gain good accuracy for fetal brain recognition.

Some other works related to fetal brain analysis were reported in [8], using machine learning algorithms and others using deep convolutional methods [35, 36] for the analysis of fetal brain abnormalities. These works help doctors, radiologists, and medical practitioners in making reliable clinical decisions and reducing false negatives. On this similar aspect, one work presented the state-of-art that had used various machine learning and deep learning works on fetal brain analysis [30].

2.2 Literatures on fetal heart

The classification of fetal heart abnormalities from ultrasound images has also received significant attention in recent years. One of the studies proposed a deep learning-based method for automatic congenital heart defects detection from fetal echocardiography [21]. In another work, the authors used a CNN architecture for the classification of the ultrasound planes [27].

A very unique and informative study is presented in [10] that presented the automatic assessment of fetal heart using the four-chamber heart ultrasound images.

Similarly, another study using ultrasound videos is presented in [25] for real-time detection of cardiac objects using the convolutional deep model. The work presented in the [5] uses ultrasound videos with deep learning algorithms for the segmentation of the ventricular septum in fetal cardiac scans using time-series information.

Besides using imaging scans, many authors have also used fetal echocardiography for the identification of fetal heart abnormalities and other parameter estimations [9, 20]. These works are using AI-based methodologies for making the judgment toward set objectives. The automation of fetal echocardiography helps in making and predicting decisions with much precision.

2.3 Literatures on other fetal organs

A few studies have also addressed the classification of other fetal organs such as lungs and kidneys. For example, a study proposed in [34] used a deep learning-based method for the automatic classification of fetal lung maturity from ultrasound images. The authors used a 2D CNN with transfer learning to predict fetal lung maturity, achieving an accuracy of 83.8% on the used dataset.

Similarly, another study presented the automation approach toward fetal head segmentation and classification using ultrasound videos [24] and assessment of the fetal geometry using the 3D ultrasound scans [19].

A few studies have also addressed the segmentation of other fetal organs such as lungs, liver, and kidneys from ultrasound images. For instance, a study proposed in [33] introduced an attention-based method for the simultaneous segmentation of the fetal lungs and heart from ultrasound images. While the authors of one paper introduced an intelligent deep learning-based framework toward segmentation of the four different parts of the ultrasound images [28].

Similarly, another study presented a comparative study between the machine learning algorithm and deep learning framework used for fetal femur length assessment [38]. This study uses the fully automated approach for making the maximum possible accurate prediction of fetal femur length using ultrasound scans.

These studies demonstrate the potential of deep learning for fetal organ classification and segmentation in ultrasound images and highlight the importance of developing accurate and reliable methods for prenatal diagnosis and treatment. Overall, these studies demonstrate the potential of deep learning for fetal organ classification and segmentation. It also highlights the importance of continued research in this area for improving prenatal diagnosis and treatment.

3 Proposed methodology

Using ultrasound imaging, the automated classification of multiple fetal organs is a challenging problem. Since the anatomy of ultrasound imaging is somewhat more complex than the existing other medical imaging modalities, thus handling the ultrasound images requires high precision and expertise. Besides many other diagnosis conditions, ultrasound images are most commonly used in maternal–fetal analysis. The whole pregnancy journey is divided into three trimesters out of which the first trimester is the highly sensitive duration as most of the organs of the baby start developing in this duration.

Overall, fetal organ analysis in the first trimester is crucial for identifying potential fetal abnormalities, assessing pregnancy risk, confirming gestational age, and monitoring fetal growth. Early detection of fetal abnormalities allows for timely medical intervention and treatment, improving the health outcomes of both the mother and the baby. Thus, in the proposed approach the objective is to classify the fetus's organs using an AI-based approach. The general flow architecture of the adopted AI pipeline is shown in Fig. 3.

3.1 Detailed description of ai pipeline

The AI framework generally includes the pipelined architecture that represents how the data flows in the system. The description of all these building blocks of the AI framework is given as:

– Data collection: Collect high-quality fetal ultrasound images from various sources. The images should cover different gestational ages and have a variety of fetal organs.

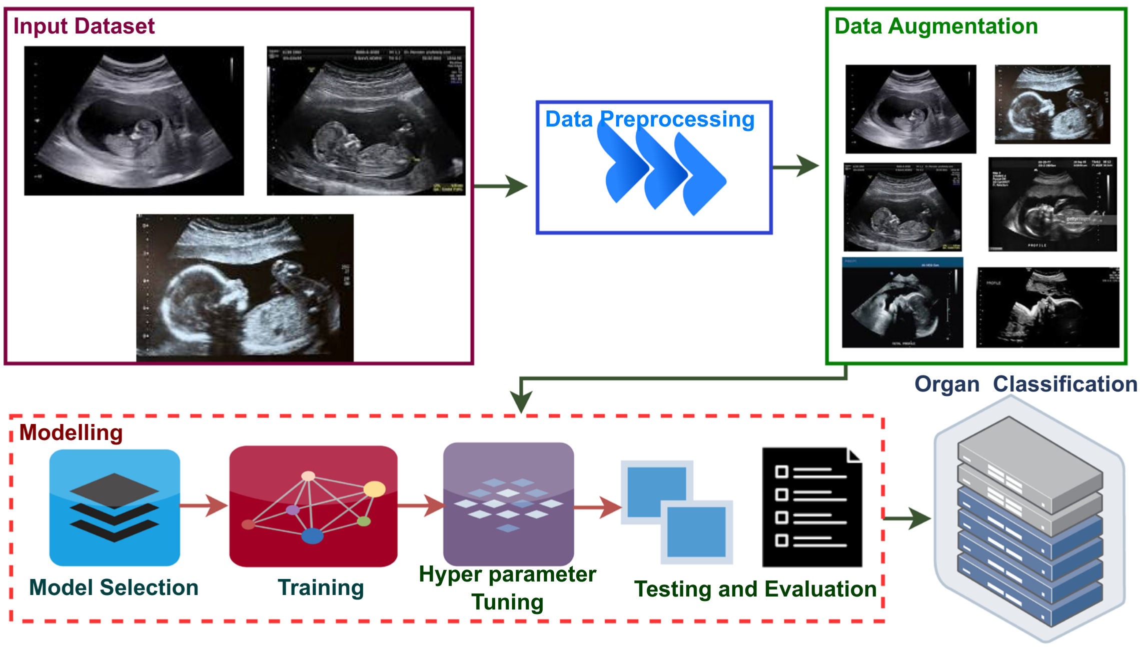

Fig. 3 AI pipeline of the proposed methodology

- Data preprocessing: Clean and preprocess the collected data by removing duplicates, resizing the images to a uniform size, and normalizing the pixel intensities.
- Data augmentation: Augment the data by applying random transformations such as rotation, flipping, and zooming to increase the size of the dataset.
- Model selection: Choose a deep learning model that is suitable for image classification tasks. Convolutional neural networks (CNNs) are widely used for image classification tasks and can be a good choice.
- Model training: Train the selected model on the preprocessed data. Use a portion of the dataset for validation during training to monitor the performance of the model.
- Hyperparameter tuning: Optimize the hyperparameters of the model, such as the learning rate, batch size, and number of epochs, to improve the model's performance.
- Testing and evaluation: Test the model on a separate dataset that was not used during training or validation. Use evaluation metrics such as accuracy, precision, and recall to evaluate the model's performance.
- Deployment: Deploy the trained model on a suitable platform such as a web application or mobile app to make it accessible to users.

3.2 Proposed dense deep neural architecture

Deep neural networks are a type of machine learning algorithm that has gained popularity in recent years due to their ability to learn complex patterns and relationships in data. Compared to traditional machine learning algorithms, deep neural networks are able to extract higher-level features from raw data, allowing for more accurate predictions and classifications. Deep neural networks can also handle a wider variety of data types, including images, speech, and text, making them highly versatile for a range of applications.

The architecture of a deep neural network consists of multiple layers of interconnected nodes, called neurons. Each neuron processes input data using a mathematical function, and the outputs from each neuron are passed to the next layer until the final output is produced. The main advantage of deep neural networks over traditional machine learning algorithms is their ability to automatically learn features from data, without the need for manual feature engineering. This means that deep neural networks can be applied to a wider range of data types and can achieve better performance than traditional machine learning algorithms.

However, training deep neural networks can be computationally expensive and requires large amounts of labeled data. Additionally, deep neural networks are often considered to be "black boxes," meaning that it can be difficult to interpret how they are making their predictions or classifications.

Here, in this work, a new deep neural architecture is proposed to classify the different organs in ultrasound images. This neural architecture is shown in Fig. 4.

The proposed dense deep neural model is a convolutional neural network (CNN) architecture that is designed

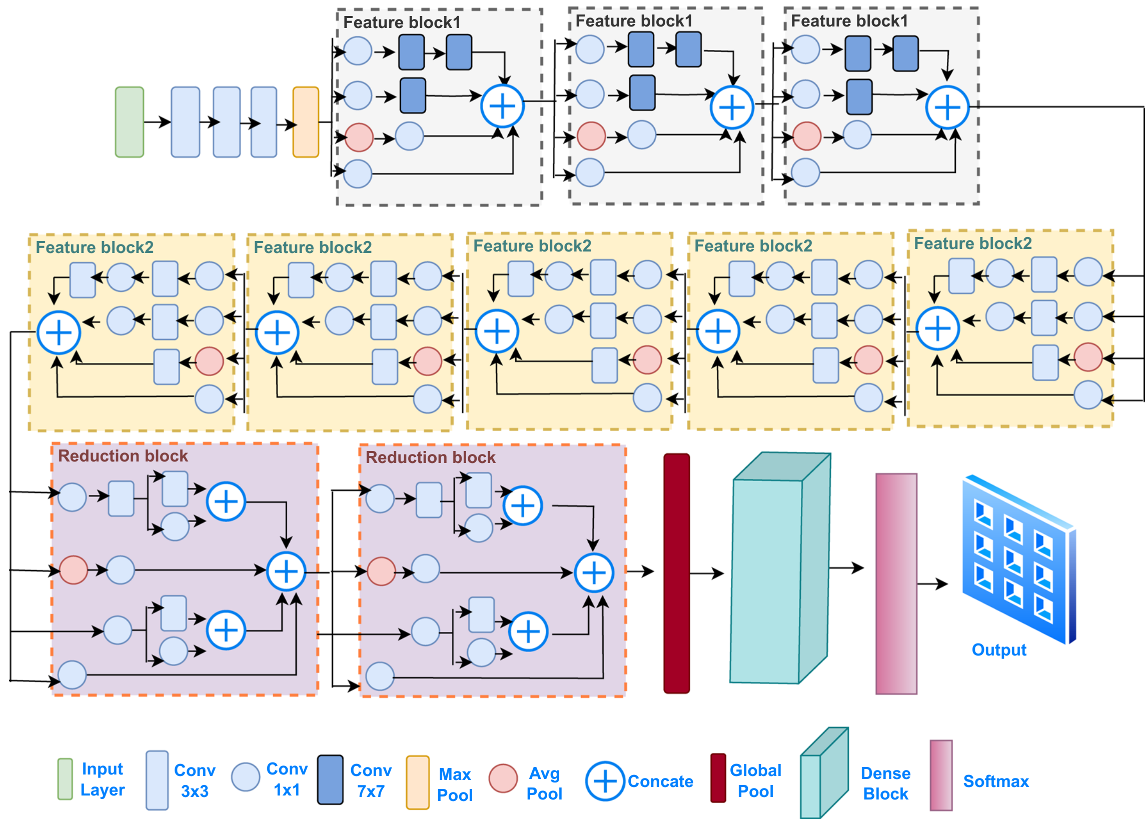

Fig. 4 Proposed deep neural architecture for multi-organ classification

to improve the computational efficiency and accuracy of the existing benchmark models. The proposed neural model uses a combination of different types of convolutional layers, including regular convolutional layers, 1x1 convolutions, and blocks of layers that use different sizes of convolutions with the concatenation layer that concatenates their outputs to capture multi-scale features.

Besides this, the model also uses *batch normalization*, which helps to accelerate the training process and improve the generalization of the model. *Pooling layers*, which reduce the spatial size of the feature maps while preserving their important features, *dense blocks*, which are added to the network at last but in between the fully connected and Feature computing layer. These dense blocks help to reduce the problem of vanishing gradients during training and improve the generalization of the model. At the last, the model also uses *Spatial averaging*, which involves averaging the output of the final layer of the network over the spatial dimensions to produce a single output value for each class.

Here is a layer-wise description of the proposed dense deep neural architecture:

1. Input layer: The first layer of the proposed model architecture is the input layer that takes an input sample, preferably an ultrasound image of the fetus, of dimension 256 $\times$ 256.
2. Cascaded block: The cascaded block of the proposed model consists of several convolutional layers with different filter sizes, followed by max pooling and batch normalization layers.
3. Multivariate feature module: The proposed model uses a series of cascaded blocks, which are having layers that use different sizes of convolutions and concatenate their outputs to capture multi-scale features. The feature module has several variations compared to the other state-of-art models, including the use of 7 $\times$ 7 convolutions, 3 $\times$ 3 convolutions, and max pooling operations.
4. Feature reduction module: The reduction module is used to reduce the spatial size of the feature maps while increasing the number of channels. The reduction module in the proposed model uses a combination of max pooling, 3 $\times$ 3 convolutions, and 1 $\times$ 1 convolutions to achieve this.
5. Dense blocks: The proposed model uses two dense blocks that are added to the network before the fully connected layer. Each dense block includes the interconnection of the layers using skip connections. It is used to reduce the problem of vanishing gradients

during training and improve the generalization of the model.

6. Global average pooling: The output of the last dense block is passed through a global average pooling layer, which produces a single feature vector for each class.
7. Softmax classifier: The final layer of the proposed model is a softmax layer that produces the probabilities for each class.

The detailed layered architecture of the proposed model with the parameter setting is shown in Table 1. This is to be noted that in the proposed design the various convolution blocks are interlinked in a cascaded manner which is not a linear pipeline. The two terms have different meanings in architectural design and represent different understandings.

The main difference between a Cascaded Deep Network and a Linear Pipeline Network lies in their architecture and information flow. The Cascaded Deep Network utilizes multiple neural networks connected in a hierarchical cascade to learn hierarchical representations, while a Linear Pipeline Network follows a linear sequence of processing stages for sequential transformations without using full neural networks at each stage.

3.3 Detailed description of the model architecture

The proposed architecture has a different architectural design than the traditional deep models which makes the model more effective in capturing complex features and provides good results. The proposed dense deep neural architecture with feature extraction and reduction blocks helps achieve good accuracy by efficiently capturing and processing informative features, learning hierarchical representations, managing computational complexity, improving generalization, supporting transfer learning, and facilitating multi-task learning. This architecture's combination of efficient feature extraction and hierarchical learning capabilities makes it a robust and versatile solution for various machine learning tasks, contributing to improved accuracy and performance. The feature extraction blocks in the architecture are designed to capture informative and discriminative features from the input data. The densely connected layers facilitate effective feature reuse, ensuring that the model can learn from a wide range of spatial and contextual information. By efficiently extracting relevant features, the model can focus on the most discriminative aspects of the data, leading to improved accuracy in the classification or regression task.

Table 1 Detailed layer-wise description of the proposed model with parameter settings

Operation	Filter size	Stride	Depth	Output
Input				256 × 256 × 3
Convolution	3×3	2	1	128 × 128 × 32
Convolution	3×3	2	1	64 × 64 × 64
Convolution	3×3	1	1	32 × 32 × 128
Max pool	3×3	2	0	32 × 32 × 256
Feature block1			3	32 × 32 × 288
Feature block2			5	16 × 16 × 768
Reduction block			2	8 × 8 × 1280
Pooling	3×3	2	0	8 × 8 × 2048
Dense block	1×1 3×3	2	3	1 × 1 × 2048
Softmax			0	1 × 1 × 1000

Also, the dense connectivity allows the model to build hierarchical representations. The architecture can effectively learn low-level features in the initial layers and then gradually capture higher-level, more abstract features in subsequent layers. This hierarchical representation learning helps the model recognize complex patterns and relationships within the data, leading to better accuracy in understanding the underlying structure of the problem. Whereas, the reduction blocks downsample the feature maps, reducing their spatial dimensions while preserving critical information. This downsampling reduces the computational load during training and inference, enabling larger batch sizes and faster processing. By managing computational complexity, the model can efficiently handle large datasets and scale well to more extensive and challenging problems.

It is also observed that the proposed dense deep neural architecture has the ability to learn hierarchical and discriminative features, along with the regularization effect of densely connected layers, which helps improve generalization. The model is less prone to overfitting and can better adapt to new, unseen data, resulting in better accuracy on test or validation datasets. Moreover, the proposed model can be effectively used for transfer learning as well. The initial layers, which capture low-level features, can be pretrained on a related task or a large dataset. Subsequent layers can then be fine-tuned on the target task, leveraging the learned representations from earlier stages. This transfer learning approach often leads to improved accuracy, especially when training data is limited.

Another key aspect of the model design is the use of feature extraction and feature reduction blocks. Although the number of individual blocks used in the proposed model varies a lot but they have a different use and the selected number of units are taken experimentally. Both these blocks in the deep neural architecture have the specific roles. Feature extraction blocks in deep learning models are responsible for learning hierarchical representations from the input data. As we move deeper into the network, the receptive field of the filters in the feature extraction blocks increases, allowing the model to capture more complex and abstract features. By having multiple feature extraction blocks, the model can build increasingly higher-level representations of the data, enabling it to understand and recognize intricate patterns and structures.

On the other hand, feature reduction blocks, often referred to as pooling or downsampling layers, help reduce the spatial dimensions of the feature maps while retaining important information. These blocks help in compressing the spatial information, which reduces computational complexity and memory requirements. However, as the information is pooled or downsampled, some fine-grained details may be lost. To address this, the architecture typically employs multiple feature extraction blocks that precede each feature reduction block. This way, the model can still capture important low-level details in earlier blocks before downsampling, while the higher-level semantics are learned in the later stages.

Generally, having more feature extraction blocks than feature reduction blocks allows the deep learning model to learn hierarchical representations and capture complex features from the input data. This design choice strikes a balance between learning high-level semantics and preserving essential low-level details. The gradual reduction of spatial dimensions in feature reduction blocks, coupled with the hierarchical learning in feature extraction blocks, enables the model to achieve good accuracy and performance in various applications.

4 Experimentation results

4.1 Dataset description

To evaluate the model performance, an open-source dataset is used having multiple organs which were annotated by experienced radiologists. This supervised dataset is openly available to the research community and is named "*FETAL-PLANES-DB: Common maternal–fetal ultrasound images*".[1]

The available dataset is having the class distribution as 6 classes, out of which the four classes are of the fetal organs (Abdomen, Brain, Femur, and Thorax), 1 class is of the mother's cervix, and 1 class is for any other organ annotated by the radiologists. Also, the brain class itself is further divided into the 3 fetal planes. But in this study, the sub-division is not considered and all three brain planes are considered as a single Brain class.

The final dataset has 12,400 images in total collected from 1792 female patients who went for an ultrasound in their first trimester. The class-wise description of the dataset is shown in Table 2.

4.2 Data augmentation

From the data analysis, it is found that there is a large variation in the sample size of the individual classes. Here, the highest number of samples is of class *Others - 4213* and the least is with class *Abdomen - 711*. Thus, there is a high chance that the model will undergo overfitting in the training phase. To overcome this issue, several data augmentation approaches were used for sample length normalization.

Data augmentation is a common technique used in machine learning to increase the size of a training dataset by generating new samples from existing ones. It can help improve the robustness of a model and prevent overfitting. Here are some common data augmentation approaches:

1. Image flipping (A1): Images can be horizontally or vertically flipped to create new samples.
2. Image rotation (A2): Images can be rotated by a certain angle to generate new samples.
3. Image cropping (A3): A portion of an image can be cropped and used as a new sample.
4. Image scaling (A4): Images can be scaled up or down to create new samples.
5. Image translation (A5): Images can be translated in a certain direction to create new samples.
6. Adding noise (A6): Adding random noise to images can create new samples.

The impact of the data augmentation on the dataset is shown in Table 3. It is to be noted that not all of the augmentation approaches were used for each class. Rather, there is a selection of the augmentation approach for respective classes. The description of this is also given in the last column of Table 3.

4.3 Dataset splitting

Dataset splitting is a common technique used in machine learning to evaluate the performance of a model. The dataset is split into two or three subsets: a training set, a validation set, and a test set.

Training set is the subset of the data used to train the model. The model learns the patterns and relationships between the input features and the target output by optimizing its parameters on this subset of the data. Whereas, *Validation set* is the subset of the data used to evaluate the performance of the model during training. The validation set helps to prevent overfitting by evaluating the model's generalization performance on unseen data. The validation set is used to tune the model's hyperparameters, such as

[1] https://zenodo.org/record/3904280.Y-xNa61BzIU.

Table 2 Description of the dataset with patient count and available sample images

Fetal organ	Patient count	Sample images
Abdomen	595	711
Brain	1080	3092
Femur	754	1040
Thorax	755	1718
Cervix	917	1626
Others	734	4213

learning rate, regularization parameters, and optimizer. Finally, the *Testing set* is the subset of the data used to evaluate the performance of the trained model after it has been fully trained. The test set is used to estimate the model's generalization performance on unseen data. The test set is typically used only once, after all, hyperparameters have been tuned, to ensure that the model's performance is not biased toward the validation set.

However, there are different ways to split the data into these subsets, such as:

1. Hold-out validation: In this method, the dataset is randomly split into training and validation sets in a predefined ratio, such as 90–10, 80–20, 70–30, 60–40, or 50–50. This splitting method completely depends on the choice and available dataset size. This method is simple to implement but can lead to high variance in the performance estimates.
2. K-fold cross-validation: In this method, the dataset is divided into k equally sized folds, where k is typically between 5 and 10. The model is trained k times, each time with a different fold used for validation and the remaining folds used for training. This method provides a more reliable estimate of the model's performance but requires more computation.
3. Stratified sampling: This method ensures that the data is split in a way that preserves the class distribution across the different subsets. This is particularly important when dealing with imbalanced datasets.

In this work, k-fold cross-validation is used with the value of $k = 5$.

4.4 Hyperparameter setting

Hyperparameter tuning is an important aspect of building effective image classification models. The hyperparameters are the configuration settings that are not learned during the training process but are set manually before training the model. The hyperparameters include the learning rate, batch size, number of epochs, regularization parameters, and optimization algorithm, among others. These parameters can significantly impact the performance of an image classification model. It can also help to achieve optimal model performance, improve generalization, and save time and resources.

The proposed model architecture thus uses certain hyperparameter optimizations whose description is given as follows:

1. Learning rate: The learning rate determines how much the weights of the neural network are updated during training. A high learning rate can result in unstable training, while a low learning rate can result in slow convergence. The learning rate can be adjusted based on the dataset and the training performance.
2. Batch size: The batch size determines how many examples are processed in each iteration during training. Larger batch sizes can speed up training but can also require more memory.

Table 3 Description of the dataset with and without augmentation

Fetal organ	Before augmentation Images count	After augmentation Images count	Augmentation method	Remark
Abdomen	711	4266	A1–A6	Whole dataset is processed with each of the augmentation method
Brain	3092	4638	A1, A3	Full dataset is processed with A1, and Randomly selected half dataset with A3
Femur	1040	4160	A2–A5	Whole dataset is processed with each of the augmentation methods A2–A5
Thorax	1718	4295	A1, A3, A6	Full dataset is processed with A1, A3 and Randomly selected half dataset with A6
Cervix	1626	4065	A2, A5, A6	Full dataset is processed with A2, A5 and Randomly selected half dataset with A6
Others	4213	4213	–	No augmentation required
Total	**12,400**	**25,637**		

3. Number of epochs: The number of epochs is the number of times the entire training dataset is processed during training. The number of epochs can be adjusted based on the convergence of the model.
4. Dropout rate: Dropout is a regularization technique that randomly drops out some of the neurons during training to prevent overfitting. The dropout rate determines the percentage of neurons that are dropped out during training.
5. Weight decay: Weight decay is a regularization technique that adds a penalty term to the loss function to prevent overfitting. The weight decay value determines the strength of the penalty term.
6. Activation function: The proposed deep model uses the Rectified Linear Unit (ReLU) activation function by default. Other activation functions such as the hyperbolic tangent (tanh) or the sigmoid function can also be used.
7. Optimizer: The optimizer determines how the weights of the neural network are updated during training. The proposed models use stochastic gradient descent (SGD) optimizer. Other optimizers such as Adam or RMSprop optimizer can also be used.

These hyperparameters can be tuned through a combination of manual tuning and automated methods such as grid search or Bayesian optimization to find the optimal values for a particular image classification task. In this work, the manual tuning of the parameters is used as the default setting. The description of the parameters with their values is shown in Table 4.

Table 4 Proposed model hyperparameters with their values and description

Hyperparameter	Values
Learning rate	0.01
Batch size	20
Epochs	100
Dropout	0.4
weight decay	0.1
Activation function	ReLU
Optimizer	SGD

4.5 Results

The implementation and experimentation of the proposed model are evaluated on the high configuration NVIDIA's DGX supercomputing Workstation. The used DGX machine is having four high-end volta graphical processing units (GPU) of Tesla architecture. Here, at each GPU we have 5000 cores which resulted in the total number of available cores as 20,000.

The experimentation result of the proposed model on the fetal organs dataset having 6 classes is shown in Table 5. The model is evaluated using the five performance metrics, i.e., accuracy, precision, recall, F-score, and AUC-ROC value. The results in Table 5 present the class-wise performance measures using the k-fold cross-validation with $k = 5$. The last row of the table presents the average value of the complete system.

The whole experimentation is performed on the epoch size of 100. It is found that the performance of the model gradually improves as the number of epochs increases. Also, it is found that the performance of the model almost converges after the epoch size of 100. The performance graph of the model using the training and validation data on epoch sizes to accuracy is presented in Fig. 5 and with loss is presented in Fig. 6.

5 Comparative analysis

The proposed model when compared with some of the benchmark deep network models is found relevant to use for the classification of the ultrasound images. In the comparative analysis, some of the models were randomly chosen from various sources. The selected models are LeNet, VGG-16, ResNet-101, ResNeXt-101, DenseNet-169, Xception, and Inception-v3. All these models are used in the form of *Transfer Learning*, where the models were pretrained on the other database and later used for the ultrasound images. The comparative analysis of the benchmark models with the proposed model is shown in Table 6.

Table 5 Experimentation result of the proposed model on the fetal organ dataset

Organs	Accuracy(%)	Precision(%)	Recall(%)	F-score(%)	AUC-ROC (%)
Abdomen	97.55	96.08	96.13	96.10	96.81
Brain	97.89	98.43	96.46	97.43	97.94
Femur	96.17	97.25	96.73	96.98	97.72
Thorax	96.64	97.89	96.02	96.94	97.08
Cervix	95.81	96.04	96.75	96.39	96.82
Others	97.04	97.03	97.87	97.44	97.25
Average	96.85	97.12	96.66	96.88	97.27

Fig. 5 Performance measure of the model with epoch sizes to accuracy and loss

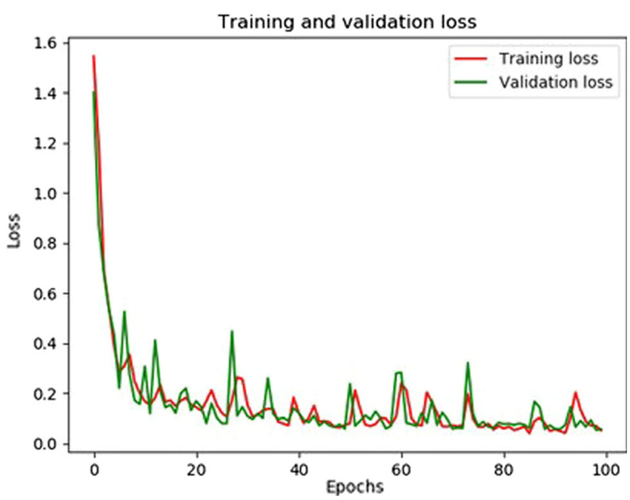

Fig. 6 Performance measure of the model with epoch sizes to accuracy and loss

In the comparison of the proposed model performance with the existing state-of-art literature, it is found that the proposed model performance really outperforms the existing works. The results of the comparison are shown in Table 7. However, it is to be noted that there exist no such works, as per the author's knowledge and search results, that have used the six classes of fetal organ planes, except one work that is reported in [1]. This work presents the comparative analysis of various existing deep models on the dataset. The performance metric used here was top errors in 1%, 3%, and accuracy. But most of the works in the literature are on using a small number of fetal planes. The comparative analysis results, as shown in Table 7, thus present the analysis of existing works with the proposed deep neural model.

6 Conclusion

This work presents the highly dense deep neural architecture for the classification of fetal organ planes using ultrasound images. The proposed deep neural architecture is designed from scratch using the core building blocks of the deep network architectures. The model is evaluated using the dataset of fetal organs having six classes, i.e., Abdomen, Brain, Thorax, Cervix, Femur, and others. The model is tested on the Nvidia DGX supercomputing platform, incorporated with the GPUs and Tensor support.

The experimentation results of the proposed model on the ultrasound dataset give best-in-class performance using various metrics such as accuracy—96.85%, precision—97.12%, recall—96.66%, F1-score—96.88%, and AUC-ROC score—97.27%. All of these experimentations were performed when the model was executed on the fixed epoch count of 100. However, when the model performance was plotted on the graph with per epoch accuracy and loss, it was found that the model gradually increased the performance as the epoch size increased. When the model tested for epoch size beyond 100, it was observed

Table 6 Comparative result of the proposed model with other Benchmark models

Organs	Accuracy (%)	Precision (%)	Recall (%)	F-score (%)	AUC-ROC (%)
LeNet	82.41	84.12	83.33	83.72	82.51
VGG-16	86.27	85.94	85.08	85.50	84.60
ResNet-101	90.09	91.28	92.14	91.70	91.45
ResNeXt-101	92.75	93.73	91.81	92.76	91.81
Inception-v3	94.82	94.51	94.61	94.55	95.08
Xception	93.48	95.22	94.72	94.96	94.92
DenseNet-169	95.49	96.37	95.87	96.11	96.76
Proposed model	96.85	97.12	96.66	96.88	97.27

The presented results are the average results computed after k-fold cross-validation with k=5. The rest of the other parameters remain constant for all models

Table 7 Comparative analysis of the proposed model with state-of-art using various parameters

Author	Fetal organ	Training set	Testing set	Performance measure
Wu et al. [32]	Abdomen	8072 2D	2606 2D	AUC = 0.98, Acc = 0.98, Rec=0.96, Spec=0.97
Cai et al. [3]	Abdomen	1292 2D	324 2D	Prec = 0.96, Rec = 0.96, F1 = 0.96
Qu et al. [23]	Braina	15314 2D	3828 2D	Acc =0.93, Prec = 0.93, Rec = 0.92, F1 = 0.93
Montero et al. [18]	Brain	6498 2D	2249 2D	Acc = 0.81, AUC = 0.86, F1 = 0.80
Kong et al. [12]	Multiple	17036 2D	5678 2D	Prec = 0.98, Rec = 0.98, F1 = 0.98
Liang et al. [14]	Multiple	17840 2D	4455 2D	Acc = 0.99, Rec= 0.96, Spec = 0.99, F1 = 0.95
Meng et al. [17]	Multiple	12000 2D	5500 2D	F1 = 0.77, Rec = 0.77, Prec = 0.78
Zhang et al. [37]	Multiple	2460 2D	820 2D	mAP = 0.95, Acc = 0.95, Prec = 0.95, Rec = 0.93
Cai et al. [2]	Multiple	280 Video	280 Video	Prec = 0.98, Rec = 0.85, F1 = 0.87
P. Sridhar et al. [27]	Multiple	3109 2D	965 2D	Acc = 0.975, Prec = 0.764%, Rec = 0.754%
Xie et al. [36]	Binary	24680 2D	4739 2D	Acc = 0.963, Prec = 0.969%, Rec = 0.959, AUC−0.989%
Lee et al. [13]	Multiple	1504 2D	752 2D	Prec= 0.75%, Rec=0.73, F1−0.74%
Burgos-Artizzu et al. [1]	Multiple	7129 2D	5271 2D	6.2% top-1 error, 0.27% top-3 error, Acc=0.94
Chen et al. [4]	Multiple	2438 videos	200 videos	Prec=0.89,Rec=0.90, F1=0.89
Tan et al. [29]	Multiple	22757 2D	5737 2D	Acc=0.70
Proposed	Multiple	5-Fold CV (25637)	5-Fold CV (5127)	Acc = 0.968, Prec = 0.97, Rec = 0.966, F1 = 0.96, AUC = 0.97

ACC accuracy, *Prec* precision, *Rec* recall, *F1* F1-score, *AUC* AUC-ROC value, Spec specificity, *mAP* mean average precision

that the performance almost became static with no further changes in accuracy level. Indeed, the validation of the model also gave the same accuracy level as we got in the training.

The overall comparison of the proposed model with the benchmark models using the same environmental conditions gives an intuition about the superiority of the proposed model to capture a better feature representation than other models. After the proposed model (Acc = 96.85%), the second most closest model is DenseNet-169 (Acc = 95.49%). Hence, there is an increment of 1.42% in the accuracy using the proposed model.

Additionally, the model is contrasted with state-of-the-art techniques; however, no such work has been discovered that uses their own constructed network, a variety of performance indicators, and a selection of organ planes to classify fetal organ planes into six classes. There are several works that used the single organ classification like Abdomen or Brain using 2D data. These works, when compared with the proposed work, found that using the limited dataset the accuracy claimed is high but when compared with the increase in the dataset size the accuracy of those models decreases. With the proposed model, the accuracy level is quite high (96.8%) even when experimented with a bigger dataset size having six classes.

On the other hand, when the proposed model is compared with other models that focused on the multiple organs classification, it is found that the proposed model performance still outperforms the other models. Indeed, the models in the literature, which were compared with the proposed model and are shown in Table 7, have slightly higher accuracy than the proposed model but all those studies have used the smaller dataset count as compared with our experimentation dataset. The overall comparison with the works gives confidence about the model's acceptability for the classification of fetal organs in ultrasound images.

In future, the model can be tested on various other application areas with the transfer learning method. Also, the work can be extended to the classification of the brain planes which was not considered in this work. A more versatile model can be designed in future that can solve both of the classification problems simultaneously, i.e., Fetal organs and Brain planes.

Funding The author declares that there is no funding associated with this project.

Data availability The dataset and code generated during and/or analyzed during the current study are available from the corresponding author at a reasonable request.

Declarations

Conflict of interest The authors of this manuscript declare that there is no conflict of interest.

Ethical approval The author of this manuscript confirms that: (i) informed written consent has been obtained from the relevant sources wherever required; (ii) all procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and its later amendments; and (iii) approval and/or informed consent was not required for the study as the dataset is collected from an open-source website which is freely available.

References

1. Burgos-Artizzu XP, Coronado-Gutiérrez D, Valenzuela-Alcaraz B, Bonet-Carne E, Eixarch E, Crispi F, Gratacós E (2020) Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. Sci Rep 10(1):1–12
2. Cai Y, Droste R, Sharma H, Chatelain P, Drukker L, Papageorghiou AT, Noble JA (2020) Spatio-temporal visual attention modelling of standard biometry plane-finding navigation. Med Image Anal 65:101762
3. Cai Y, Sharma H, Chatelain P, Noble JA (2018) Multi-task sonoeyenet: detection of fetal standardized planes assisted by generated sonographer attention maps. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I, pp. 871–879. Springer
4. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D (2019) Self-supervised learning for medical image analysis using image context restoration. Med Image Anal 58:101539
5. Dozen A, Komatsu M, Sakai A, Komatsu R, Shozu K, Machino H, Yasutomi S, Arakaki T, Asada K, Kaneko S et al (2020) Image segmentation of the ventricular septum in fetal cardiac ultrasound videos based on deep learning using time-series information. Biomolecules 10(11):1526
6. Farsana WF, Kowsalya N (2022) Dilated multi-scale-link net with merged self attention based fetal head segmentation using 2d ultra sound image. Specialusis Ugdymas 1(43):8825–8839
7. Fiorentino MC, Villani FP, Di Cosmo M, Frontoni E, Moccia S (2022) A review on deep-learning algorithms for fetal ultrasound-image analysis. Med Image Anal 83:102629
8. Gofer S, Haik O, Bardin R, Gilboa Y, Perlman S (2022) Machine learning algorithms for classification of first-trimester fetal brain ultrasound images. J Ultrasound Med 41(7):1773–1779
9. Gong Y, Zhang Y, Zhu H, Lv J, Cheng Q, Zhang H, He Y, Wang S (2019) Fetal congenital heart disease echocardiogram screening based on dgacnn: adversarial one-class classification combined with video transfer learning. IEEE Trans Med Imag 39(4):1206–1222
10. Gudigar A, U R, Samanth J, Vasudeva A, Nayak K, Tan RS, Ciaccio EJ, Ooi CP, Barua PD, Molinari F et al (2022) Role of four-chamber heart ultrasound images in automatic assessment of fetal heart: A systematic understanding. In: Informatics, vol. 9, p. 34. MDPI
11. Hou B, Khanal B, Alansary A, McDonagh S, Davidson A, Rutherford M, Hajnal JV, Rueckert D, Glocker B, Kainz B (2018) 3-d reconstruction in canonical co-ordinate space from arbitrarily oriented 2-d images. IEEE Trans Med Imag 37(8):1737–1750
12. Kong P, Ni D, Chen S, Li S, Wang T, Lei B (2018) Automatic and efficient standard plane recognition in fetal ultrasound images via multi-scale dense networks. In: Data Driven Treatment Response Assessment and Preterm, Perinatal, and Paediatric Image Analysis: First International Workshop, DATRA 2018 and Third International Workshop, PIPPI 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 16, 2018, Proceedings 3, pp. 160–168. Springer
13. Lee LH, Gao Y, Noble JA (2021) Principled ultrasound data augmentation for classification of standard planes. In: International Conference on Information Processing in Medical Imaging, pp. 729–741. Springer
14. Liang J, Huang R, Kong P, Li S, Wang T, Lei B (2019) Sprnet: Automatic fetal standard plane recognition network for ultrasound images. In: Smart Ultrasound Imaging and Perinatal, Preterm and Paediatric Image Analysis: First International Workshop, SUSI 2019, and 4th International Workshop, PIPPI 2019, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings 4, pp. 38–46. Springer
15. Lin Z, Li S, Ni D, Liao Y, Wen H, Du J, Chen S, Wang T, Lei B (2019) Multi-task learning for quality assessment of fetal head ultrasound images. Med Image Anal 58:101548
16. Liu S, Wang Y, Yang X, Lei B, Liu L, Li SX, Ni D, Wang T (2019) Deep learning in medical ultrasound analysis: a review. Engineering 5(2):261–275
17. Meng Q, Rueckert D, Kainz B (2020) Unsupervised cross-domain image classification by distance metric guided feature alignment. In: Medical Ultrasound, and Preterm, Perinatal and Paediatric Image Analysis: First International Workshop, ASMUS 2020, and 5th International Workshop, PIPPI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4-8, 2020, Proceedings 1, pp. 146–157. Springer
18. Montero A, Bonet-Carne E, Burgos-Artizzu XP (2021) Generative adversarial networks to improve fetal brain fine-grained plane classification. Sensors 21(23):7975

19. Oghli MG, Shabanzadeh A, Moradi S, Sirjani N, Gerami R, Ghaderi P, Taheri MS, Shiri I, Arabi H, Zaidi H (2021) Automatic fetal biometry prediction using a novel deep convolutional network architecture. Phys Med 88:127–137
20. Pu B, Zhu N, Li K, Li S (2021) Fetal cardiac cycle detection in multi-resource echocardiograms using hybrid classification framework. Future Gener Comput Syst 115:825–836
21. Qiao S, Pang S, Luo G, Pan S, Yu Z, Chen T, Lv Z (2022) Rlds: an explainable residual learning diagnosis system for fetal congenital heart disease. Future Gener Comput Syst 128:205–218
22. Qu R, Xu G, Ding C, Jia W, Sun M (2019) Deep learning-based methodology for recognition of fetal brain standard scan planes in 2d ultrasound images. IEEE Access 8:44443–44451
23. Qu R, Xu G, Ding C, Jia W, Sun M (2020) Standard plane identification in fetal brain ultrasound scans using a differential convolutional neural network. IEEE Access 8:83821–83830
24. Rasheed K, Junejo F, Malik A, Saqib M (2021) Automated fetal head classification and segmentation using ultrasound video. IEEE Access 9:160249–160267
25. Sapitri AI, Nurmaini S, Rachmatullah MN, Tutuko B, Darmawahyuni A, Firdaus F, Rini DP, Islami A (2023) Deep learning-based real time detection for cardiac objects with fetal ultrasound video. Inf Med Unlocked 36:101150
26. Shen YT, Chen L, Yue WW, Xu HX (2021) Artificial intelligence in ultrasound. Eur J Radiol 139:109717
27. Sridar P, Kumar A, Quinton A, Nanan R, Kim J, Krishnakumar R (2019) Decision fusion-based fetal ultrasound image plane classification using convolutional neural networks. Ultrasound Med Biol 45(5):1259–1273
28. Sun Y, Yang H, Zhou J, Wang Y (2022) Issmf: integrated semantic and spatial information of multi-level features for automatic segmentation in prenatal ultrasound images. Artif Intell Med 125:102254
29. Tan J, Au A, Meng Q, Kainz, B (2019) Semi-supervised learning of fetal anatomy from ultrasound. In: Domain Adaptation and Representation Transfer and Medical Image Learning with Less Labels and Imperfect Data: First MICCAI Workshop, DART 2019, and First International Workshop, MIL3ID 2019, Shenzhen, Held in Conjunction with MICCAI 2019, Shenzhen, China, October 13 and 17, 2019, Proceedings 1, pp. 157–164. Springer
30. Torres HR, Morais P, Oliveira B, Birdir C, Rüdiger M, Fonseca JC, Vilaça JL (2022) A review of image processing methods for fetal head and brain analysis in ultrasound images. Comput Methods Progr Biomed 215:106629
31. Wang L, Nie D, Li G, Puybareau É, Dolz J, Zhang Q, Wang F, Xia J, Wu Z, Chen JW et al (2019) Benchmark on automatic six-month-old infant brain segmentation algorithms: the iseg-2017 challenge. IEEE Trans Med Imag 38(9):2219–2230
32. Wu L, Cheng JZ, Li S, Lei B, Wang T, Ni D (2017) Fuiqa: fetal ultrasound image quality assessment with deep convolutional networks. IEEE Trans Cybernet 47(5):1336–1349
33. Xi J, Chen J, Wang Z, Ta D, Lu B, Deng X, Li X, Huang Q (2021) Simultaneous segmentation of fetal hearts and lungs for medical ultrasound images via an efficient multi-scale model integrated with attention mechanism. Ultrason Imag 43(6):308–319
34. Xia TH, Tan M, Li JH, Wang JJ, Wu QQ, Kong DX (2021) Establish a normal fetal lung gestational age grading model and explore the potential value of deep learning algorithms in fetal lung maturity evaluation. Chin Med J 134(15):1828–1837
35. Xie B, Lei T, Wang N, Cai H, Xian J, He M, Zhang L, Xie H (2020) Computer-aided diagnosis for fetal brain ultrasound images using deep convolutional neural networks. Int J Comput Assis Radiolo Surg 15:1303–1312
36. Xie H, Wang N, He M, Zhang L, Cai H, Xian J, Lin M, Zheng J, Yang Y (2020) Using deep-learning algorithms to classify fetal brain ultrasound images as normal or abnormal. Ultrasound Obstetr Gynecol 56(4):579–587
37. Zhang B, Liu H, Luo H, Li K (2021) Automatic quality assessment for 2d fetal sonographic standard plane based on multitask learning. Medicine 100(4):e24427
38. Zhu F, Liu M, Wang F, Qiu D, Li R, Dai C (2021) Automatic measurement of fetal femur length in ultrasound images: a comparison of random forest regression model and segnet. Math Biosci Eng 18(6):7790–7805