Journal of Disability Research

2024 | Volume 3 | Pages: 1–16 | e-location ID: e20240005
DOI: 10.57197/JDR-2024-0005

Enhancing Fetal Medical Image Analysis through Attention-guided Convolution: A Comparative Study with Established Models

Muna Al-Razgan[1,2], Yasser A. Ali[2,3] and Emad Mahrous Awwad[4,*]

[1]Department of Software Engineering, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[2]King Salman Center for Disability Research, Riyadh, Saudi Arabia
[3]Department of Information Systems, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[4]Electrical Engineering Department, College of Engineering, King Saud University, Riyadh 11421, Saudi Arabia

Correspondence to:
Emad Mahrous Awwad*, e-mail: 442106835@student.ksu.edu.sa, Tel: +966562939242

Received: September 6 2023; Revised: January 8 2024; Accepted: January 10 2024; Published Online: February 22 2024

ABSTRACT

The ability to detect and track fetal growth is greatly aided by medical image analysis, which plays a crucial role in parental care. This study introduces an attention-guided convolutional neural network (AG-CNN) for maternal–fetal ultrasound image analysis, comparing its performance with that of established models (DenseNet 169, ResNet50, and VGG16). AG-CNN, featuring attention mechanisms, demonstrates superior results with a training accuracy of 0.95 and a testing accuracy of 0.94. Comparative analysis reveals AG-CNN's outperformance against alternative models, with testing accuracies for DenseNet 169 at 0.90, ResNet50 at 0.88, and VGG16 at 0.86. These findings underscore the effectiveness of AG-CNN in fetal image analysis, emphasising the role of attention mechanisms in enhancing model performance. The study's results contribute to advancing the field of obstetric ultrasound imaging by introducing a novel model with improved accuracy, demonstrating its potential for enhancing diagnostic capabilities in maternal–fetal healthcare.

KEYWORDS

fetal medical imaging, attention-guided convolution, adaptive feature extraction, medical image analysis, convolutional neural network, diagnostic precision

INTRODUCTION

The field of fetal medical image analysis has gained significant importance due to its vital role in maternal–fetal healthcare. Accurate and efficient analysis of ultrasound images is crucial for the early detection of anomalies and ensuring the well-being of both the mother and the fetus. In this context, the present study aims to contribute to the advancement of this field by introducing a novel attention-guided convolutional neural network (AG-CNN) for enhanced feature extraction in maternal–fetal ultrasound images. The field of medical image analysis has made incredible strides in recent years, completely altering how specialists diagnose and treat patients in a wide range of fields (Horgan et al., 2023). Medical imaging of the fetus is an important part of this field as it can reveal important information about the fetus's growth and health (Mehrdad et al., 2021).

In recent years, medical imaging has played a pivotal role in diagnosing anomalies and evaluating congenital and acquired disabilities. One area of significant focus is fetal medical image analysis, where advanced techniques contribute to the early detection of abnormalities, thereby facilitating timely interventions and improved outcomes. This study delves into the realm of attention-guided convolution, presenting a novel technique for adaptive feature extraction in fetal medical image analysis. By harnessing the power of attention mechanisms, our approach aims to enhance the interpretation of ultrasound images, particularly in the context of anomalies and conditions associated with congenital and acquired disabilities. The utilisation of a large dataset from real clinical settings allows us to explore the efficacy of the proposed technique across diverse cases, ranging from morphological anomalies to neurodevelopmental conditions. This research holds promise in advancing the field, contributing valuable insights for improved diagnostic capabilities and patient care in the context of anomalies and various fetal conditions. Medical imaging of the fetus, obtained by methods such as ultrasound and magnetic resonance imaging, provides insight into the complex stages of prenatal development, allowing for the early diagnosis of anomalies and developmental diseases (Tenajas et al., 2023). While considerable progress has been made in the application of artificial intelligence (AI) to ultrasound analysis, there exists a knowledge gap in the optimal utilisation of attention mechanisms for improved feature extraction. Existing models may not fully exploit the intricate details within ultrasound images, potentially hindering diagnostic accuracy. Addressing this gap, our study seeks to explore and leverage attention-guided convolution, providing a nuanced understanding of its impact on feature extraction in fetal medical image analysis. By addressing this specific void in current knowledge, our research aims to contribute valuable insights that can refine existing methodologies and pave the way for more accurate diagnostic tools.

However, due to the specific constraints provided by fetal medical imaging, information extraction from these images remains a complicated endeavour (Xiao et al., 2023). Due to factors such as background noise, anatomical heterogeneity, and fetal position changes, the collected pictures may lack clarity and consistency (Fiorentino et al., 2022). These complexities are typically beyond the capabilities of conventional image analysis methods, calling for more cutting-edge computational ways to tackle the problems (Iskandar et al., 2023). As a result of their remarkable performance in tasks including image segmentation, classification, and feature extraction, convolutional neural networks (CNNs) have become a useful tool in medical image analysis. In contrast to manually crafted features, which may not be able to capture the intricacies found in fetal medical imaging, these deep learning models may automatically learn important features from data (Cai et al., 2018). The complexity and subtlety of fetal anatomy and development make it difficult for CNNs to be applied to fetal medical imaging, notwithstanding their effectiveness (Fergus et al., 2021).

The research problem at the core of this study revolves around the need for a more nuanced and effective approach to feature extraction in maternal–fetal ultrasound images. The proposed AG-CNN model incorporates attention mechanisms to selectively focus on relevant image regions, potentially improving diagnostic accuracy. The main challenge is coming up with an adaptive feature extraction method that can reliably detect subtle details in fetal medical photos in the presence of background noise and natural variation. When extracting features from an image, conventional CNN architectures average over the entire image, which can lead to an exaggeration of background noise and a blurring of critical anatomical details. Therefore, it is crucial to create a method that can dynamically zero down on important parts of a picture while discarding the rest. Adaptive feature extraction in fetal medical picture analysis faces a number of obstacles. There is substantial variation in the appearance of fetal structures because, first, there is no standardised fetal anatomy across different imaging sessions and gestational ages. Creating a standard method for feature extraction is made more difficult by this diversity. Second, because of the complexity and vulnerability of fetal organs, an approach that is sensitive to changes in a little amount of data while being robust to noise is desirable. Third, in medical applications, interpretability of the feature extraction process is critical so that physicians can comprehend the reasoning behind a model's predictions. To aid in tasks like organ segmentation, anomaly identification, and age estimate, fetal medical image analysis seeks to extract relevant and meaningful features from fetal images. Traditional feature extraction approaches face difficulties due to noise, anatomical heterogeneity, and fetal position changes. In order to overcome these obstacles, we present a new method dubbed "attention-guided convolution," which incorporates attention processes into the CNN architecture.

This attention-guided convolution process selectively emphasises key features while downplaying those that are less crucial to the overall meaning of the input image. The resulting characteristics should be more discriminative and immune to noise, leading to better results when analysing fetal medical images. The issue can be stated in terms of improving the feature extraction process in fetal medical picture analysis by including attention-guided convolution into the CNN architecture. The goal of this method is to enhance the precision and reliability of the analysis by compensating for noise, anatomical variability, and fetal position changes. Innovative solutions that go beyond standard CNN designs are needed to tackle these problems. In light of these challenges, this study presents "attention-guided convolution," a unique technique that combines attention processes into CNNs to improve the adaptive feature extraction process in fetal medical picture analysis. The goal of this method is to help physicians make better judgements based on the model's predictions by supplying them with interpretable attention maps.

The study's main objectives are:

- To bring the idea of attention-guided convolution to the field of fetal medical picture analysis and put it into practice. To do this, a novel method must be developed that dynamically adjusts the feature extraction process to zero in on important parts of the photos while filtering out noise and extraneous features. Our goal is to improve the CNN's ability to recognise fetal anatomy by embedding attention processes into the network's architecture.
- To enhance the process of extracting fetal anatomy from medical imaging. Our goal is to show that, despite noise, variability, and changes in the fetal position, the attention-guided convolution technique can still identify and highlight the delicate and complicated fetal organs. Our goal in doing so is to improve the precision of organ segmentation, a crucial process in the processing of fetal medical images.
- To demonstrate that the attention-guided convolution method is capable of effectively mitigating the effects of background noise and other distracting features in fetal medical photos. Our goal is to demonstrate the method's reliability across a range of conditions by conducting experimental validation across a spectrum of imaging modalities and gestational ages.
- To shed light on the working and effects of the attention mechanism on feature extraction. We hope that by producing interpretable attention maps, we may help doctors and academics better understand the model's decision-making

process, which will increase confidence in the model and speed up its adoption in clinical practice.

Several new insights into fetal medical image analysis are provided by this study. In particular, we provide a novel approach termed "attention-guided convolution," which embeds attention mechanisms directly into the framework of CNNs. To improve the accuracy and robustness of organ segmentation, anomaly detection, and age estimate, this method dynamically emphasises essential regions in fetal medical images while suppressing noise and unnecessary features. Additionally, our work generates interpretable attention maps, which provide physicians and researchers with insights into the model's decision-making process, thus overcoming the interpretability barrier. We illustrate the versatility of the attention-guided convolution method across a variety of datasets and clinical contexts, and show that it outperforms more conventional CNN methods through extensive experimental validation. Collectively, our work paves the way for better application of fetal medical pictures in clinical practice and research, leading to improved diagnostics, better understanding of fetal development, and ultimately better outcomes.

RELATED WORK

The study by Horgan et al. (2023) provides a high-level overview of AI's potential uses in obstetric ultrasound. This review takes a broad look at the present-day application of AI in the area, including its effects on diagnostic precision, workflow optimisation, and the difficulties of deploying AI technologies. Medical treatments, such as surgery and image-guided interventions, are the primary emphasis of the review of sophisticated medical telerobots provided by Mehrdad et al. (2021). The paper examines numerous facets of telerobotic systems, illuminating their successes, failures, and promising future prospects. Recent developments in ultrasound scanning with the help of AI are presented by Tenajas et al. (2023). This study explores the use of AI approaches in ultrasound systems to boost image quality, give operators immediate feedback, and streamline the scanning process.

Xiao et al. (2023) assess the state of AI in fetal ultrasonography and its potential future developments.

The authors highlight the contributions and problems of using AI approaches for diverse fetal ultrasound analysis tasks like image segmentation, anomaly identification, and gestational age calculation. The deep learning techniques for fetal ultrasound image processing are reviewed in detail by Fiorentino et al. (2022). The research addresses a wide variety of uses, such as image segmentation, classification, and detection, demonstrating the efficacy of deep learning algorithms in dealing with the intricacies of fetal ultrasound images. A study on synthesising realistic ultrasound images of the fetal brain is presented by Iskandar et al. (2023). In order to provide deep learning models with more realistic data for training, the authors propose a way to synthetically generate ultrasound images of fetal brains. In this study, we explore the feasibility of using this method to broaden the applicability of algorithms used in fetal brain analysis.

Standardised fetal ultrasound plane detection using eye tracking is presented by Cai et al. (2018). In order to improve the precision and reliability of plane localisation in clinical practice, the authors employ eye-tracking data to direct the detection of fetal ultrasound planes.

The work of Fergus et al. (2021) centres on the application of one-dimensional CNNs to the modelling of cardiotocography time-series signals that have been segmented. This study highlights the use of deep learning models trained on cardiotocography data for the early detection of aberrant delivery outcomes, demonstrating the potential of CNNs in enhancing prenatal care. A study on how to learn the architectures of deep neural networks using differential evolution is presented by Belciug (2022). The author uses this strategy to medical image processing to demonstrate the utility of evolutionary algorithms in enhancing the efficiency of neural network topologies. Automatic fetal abdominal segmentation from ultrasound pictures is proposed by Ravishankar et al. (2016) using a hybrid method. The authors show the promise of hybrid solutions for difficult segmentation tasks by accurately segmenting fetal abdominal tissues using a combination of deep learning and contour-based methods.

For fetal ultrasound picture segmentation, Zeng et al. (2021) present a deeply supervised attention-gated V-Net. The accuracy of head circumference biometry from ultrasound pictures is increased by the network design shown in this paper, which blends attention processes with segmentation models. An approach for detecting fetal movement and recognising anatomical planes is presented by Dandıl et al. (2021), which makes use of the YOLOv5 network. In order to aid in thorough evaluations of fetal health, the authors employ this network to recognise anatomical features and track fetal movement in ultrasound images. In order to automatically classify common maternal–fetal ultrasound planes, Burgos-Artizzu et al. (2020). assess deep CNNs. The authors investigate the usefulness of CNNs in the classification of maternal–fetal ultrasound images, expanding our knowledge of automatic plane recognition.

Automatic classification of frequent maternal–fetal ultrasound planes using deep CNNs is discussed and evaluated by Burgos-Artizzu et al. (2020). The research makes a contribution to the area by evaluating CNNs' ability to identify maternal–fetal ultrasound planes. For the purpose of fetal head analysis, Alzubaidi et al. (2022) offer a transfer learning ensemble method. Using transfer learning and ensemble approaches, the authors propose a comprehensive solution for multi-task analysis, in this case predicting the gestational age and weight of a fetus from ultrasound scans. Sengan et al. (2022) use deep learning to segment echocardiographic images for prenatal diagnosis of fetal cardiac rhabdomyoma. The authors' goal is to aid in the early detection of cardiac problems by using deep learning algorithms to segment photos of the fetal heart. Categorisation of Down syndrome markers using dense neural networks in fetal ultrasound pictures is presented by Pregitha et al. (2022). The use of deep neural networks to detect Down syndrome in fetal ultrasound images is investigated. For recognising standard scan planes of the fetal brain in 2D ultrasound pictures, Qu et al. (2020) present a deep learning-based solution. The authors present a system that uses deep CNNs to automatically recognise common ultrasound planes used for fetal brain scans.

Automatic classification of frequent maternal–fetal ultrasound planes using deep CNNs was evaluated by Cerrolaza et al. (2018). The authors hope that their work will help advance automated plane recognition by expanding our knowledge of CNNs' capability in recognising planes in maternal–fetal ultrasound. Deep learning methods for ultrasound during pregnancy are discussed by Diniz et al. (2021). This paper provides a survey of recent work that has used deep learning techniques to assess ultrasound images for signs of pregnancy. The study by Wang et al. (2021) provides an extensive literature review on the application of deep learning to the processing of medical ultrasound images. The authors address the influence of deep learning approaches on bettering diagnostic accuracy and clinical decision-making across a variety of settings. In their paper, Lipa and Trzciński (Płotka et al., 2022) discuss the results of a study in which deep learning fetal ultrasound video models were compared to human observers. In this work, the authors investigate deep learning models for fetal ultrasound video interpretation with the end goal of reaching the same biometric measurement accuracy as human observers. Automatic fetal biometry prediction using a unique deep convolutional network architecture is proposed by Ghelich Oghli et al. (2023). Using convolutional networks as an example, the authors present a deep learning strategy for predicting fetal biometric data. Deep learning and the Industrial Internet of Things (IIoT) are the foundation for automatic fetal ultrasound standard plane detection, which is the focus of Pu et al. (2021). In order to accurately recognise common fetal ultrasound planes, the authors offer a solution that blends deep learning approaches with IIoT concepts. Table 1 show the summarization of the related work.

An integrated method that leverages the best features of various segmentation architectures, attention mechanisms, and fusion approaches is a key area of unexplored study in the field of medical image segmentation. While many separate studies have made important contributions, there has been a dearth of meta-analyses that examine how these advances interact with one another. In addition, there is a lack of comprehensive and generalisable solutions because most studies have only looked at one or two imaging modalities or health issues. To fill this void, we need a standardised framework for medical picture segmentation that makes use of attention-guided convolution, multi-modal fusion, and adaptive architectures. In order to overcome the obstacles presented by fetal images, previous research in medical image processing has mostly focused on modifying pre-existing CNN structures. Several methods have been investigated to address the problem of insufficient training data, including transfer learning from more general medical imaging domains, domain adaptation to account for anatomical heterogeneity, and data augmentation techniques. There has been a rise in interest in the application of interpretable AI methods in the field of medical imaging. To better

Table 1: Comparative table.

Reference	Dataset	Techniques	Outcome	Limitations
Iskandar et al. (2023)	Fetus dataset	Image synthesis	Proposal of method for realistic ultrasound fetal brain imaging synthesis	No real dataset, focus on image synthesis
Cai et al. (2018)	Fetal ultrasound data, eye-tracking data	Attention mechanisms	SonoEyeNet for standardised fetal ultrasound plane detection	Limited dataset and potential hardware dependency
Fergus et al. (2021)	Cardiotocography data	1D CNNs	Modelling segmented cardiotocography time-series signals	Focus on time-series data, no ultrasound
Belciug (2022)	Fetus dataset	Differential evolution	Learning deep neural network architectures for medical imaging	No specific dataset mentioned
Ravishankar et al. (2016)	Fetal ultrasound data	Hybrid approach	Automatic segmentation of fetal abdomen	No comprehensive dataset mentioned
Zeng et al. (2021)	Fetal ultrasound data	Attention-gated V-Net	Head circumference biometry using deep learning	Limited detail on other techniques
Dandıl et al. (2021)	Ultrasound scans	YOLOv5 network	Fetal movement detection, anatomical plane recognition	Limited scope, YOLOv5 specific
Burgos-Artizzu et al. (2020)	Maternal–fetal ultrasound images	Deep CNNs	Automatic classification of maternal–fetal ultrasound planes	Focus on plane classification, no fetal outcome
Burgos-Artizzu et al. (2020)	Maternal–fetal ultrasound images	Deep CNNs	Automatic classification of maternal–fetal ultrasound planes	Similar to Burgos-Artizzu et al. (2020)
Alzubaidi et al. (2022)	Fetal head ultrasound data	Ensemble transfer learning	Fetal head analysis, gestational age, and weight prediction	Specific focus on fetal head analysis
Sengan et al. (2022)	Fetal cardiac ultrasound images	Deep learning	Echocardiographic image segmentation for diagnosing fetal cardiac rhabdomyoma	Specific focus on cardiac analysis
Pregitha et al. (2022)	Ultrasound fetal images	Dense neural network	Down syndrome marker classification	Specific focus on Down syndrome markers
Qu et al. (2020)	2D ultrasound images	Deep learning	Recognition of fetal brain standard scan planes	Specific focus on brain scan plane recognition
Cerrolaza et al. (2018)	Maternal–fetal ultrasound images	Deep CNNs	Automatic classification of maternal–fetal ultrasound planes	Similar to Burgos-Artizzu et al. (2020)
Diniz et al. (2021)	Ultrasound scans	Deep learning	Deep learning strategies for ultrasound in pregnancy	Broad review without specific dataset/technique

Abbreviation: CNN, convolutional neural network.

comprehend which aspects of an image contribute to a model's conclusion, the latter can use attention mechanisms to provide different amounts of importance to distinct regions. However, the development of attention mechanisms that are optimised for the nuances of fetal medical imagery is still in its infancy.

MATERIALS AND METHODS

The approach used in this research makes use of an attention-guided CNN model to examine a large dataset of maternal–fetal ultrasound pictures from an actual clinical scenario at BCNatal, which includes Hospital Clinic and Hospital Sant Joan de Deu in Barcelona, Spain. The dataset was painstakingly curated, and it included over 12,000 photos from 1792 individuals who were receiving standard tests in their second or third trimesters of pregnancy. The gestational age range was 18-40 weeks due to the exclusion of multiple pregnancies, congenital abnormalities, and aneuploidies. A senior maternal–fetal doctor painstakingly annotated each image in the dataset with anatomical plane labels. There are six distinct categories in the dataset, including five primary maternal–fetal anatomical planes. Multiple operators used a variety of ultrasound equipment, including the Voluson E6, Voluson S10, and Aloka systems, to gather the ultrasound images. In Figure 1, the suggested process flow is provided.

Mathematical formulation

Let's denote a fetal medical image as I, which is a two-dimensional matrix representing the pixel intensities. Our aim is to learn a set of features, F, that capture relevant anatomical structures while minimising the impact of noise and variability. Conventionally, a CNN extracts features using convolutional layers, which can be represented as:

$$F_{ij} = \sum_m \sum_n I_{(i+m)(j+n)} \cdot K_{mn},$$

where

F_{ij} is the value of the feature map at position (i, j).

$I_{(i+m)(j+n)}$ represents the pixel intensity at position $(i+m, j+n)$ in the input image.

K_{mn} is the convolution kernel applied at position (m,n).

However, conventional convolution averages over the entire image, which may enhance noise and dilute crucial details. We develop an attention mechanism that makes real-time adjustments to the convolution process to solve this problem. The attention mechanism prioritises certain parts of the image above others based on their importance to the mission at hand. The following is how we calculate the attention map, A:

$$A_{ij} = \sigma(Wa \cdot [F_{ij} \mid I_{ij}]),$$

where

A_{ij} is the attention weight at position (i, j).

Wa is the learnable attention parameter.

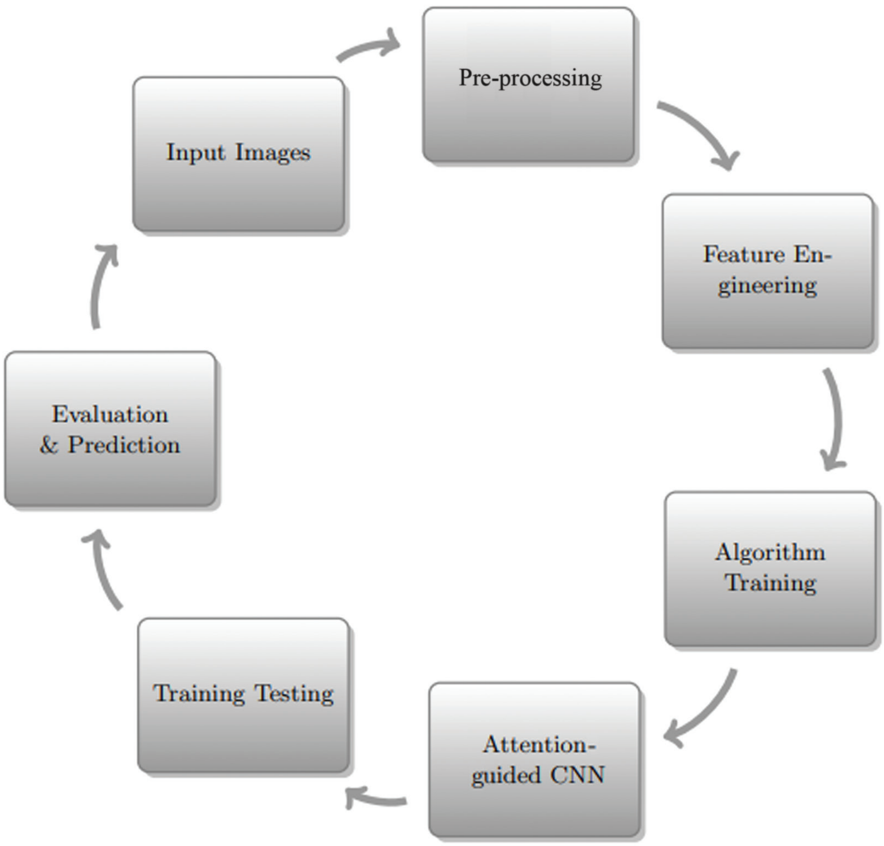


Figure 1: The proposed working flow.

σ is the activation function.

‖ denotes concatenation.

The attention-guided convolution is then formulated as follows:

$$F_{ij}^{att} = \sum_m \sum_n A_{(i+m)(j+n)} \cdot I_{(i+m)(j+n)} \cdot K_{mn},$$

where

F_{ij}^{att} is the value of the attention-guided feature map at position (i, j).

$A_{(i+m)(j+n)}$ is the attention weight at position $(i+m, j+n)$.

Dataset description

This research made use of the BCNatal dataset, which was painstakingly assembled from ultrasound photos of mothers and their unborn children taken at Barcelona's Hospital Clinic and Hospital Sant Joan de Deu. The dataset's composition, labelling, and distribution are all described in this section.

Dataset composition

More than 12,000 ultrasound scans from 1792 patients are included in the dataset's complete collection. During their second and third trimesters of pregnancy, these women went in for routine checkups. The dataset was created to be broadly representative of maternal–fetal anatomical planes, allowing for a wide range of research and potential clinical applications.

Labelling and categories

An experienced maternal–fetal doctor painstakingly assigned anatomical plane labels to each photograph in the dataset. The naming system includes not only the five most common maternal–fetal anatomical planes but also a sixth grouping for all other variations. Labels were placed on the following anatomical planes:

Fetal abdomen

This section focuses on fetal weight and fetal abdominal shape.

Fetal brain

The study of neurodevelopment benefits from labelled photos from this category.

Trans-thalamic pictures are essential for studying neurodevelopment.

Fetal weight analysis can benefit from images in the Trans-cerebellum category.

Images that help analyse the growth of the heart and lungs fall under the category of Trans-ventricular.

Fetal femur

These photos are helpful in determining the approximate birth weight of the fetus.

Images in the Fetal Thorax category help researchers learn more about how the fetal heart and lungs form.

Maternal cervix

Premature birth is studied using images from this category.

Other

This section includes a wide range of medical pictures used for a variety of applications.

The variety of ultrasound pictures obtained across multiple anatomical planes is illustrated in Figure 2 (sample photos from the collection). The intricacy and variety of the data are illustrated by the subfigures (a) through (h), which show instances of images belonging to different anatomical categories.

Dataset distribution

Table 2 displays the dataset distribution across several anatomical categories and planes. For each anatomical plane category, this table reveals the clinical relevance, patient count, and image count. There is a wide variety of clinical settings and uses represented in the collection.

The distribution of these pictures across different anatomical planes is shown graphically in Figure 3. Different anatomical categories are represented in the sub-figures, showing how each group adds variety to the dataset as a whole. With this representation, users may quickly grasp the dataset's structure and clinical relevance.

The Voluson E6, the Voluson S10, and the Aloka systems account for the bulk of the ultrasound machines represented in the dataset. Table 3 details the image distribution across these machines and their respective operators. This table shows how various machines and operators have contributed to the dataset.

Figure 4 is a visual representation of how commonly used ultrasound equipment produces particular types of images. The contribution of various machine types to the dataset is graphically represented by the bar graph. This visualisation helps to shed light on the ways in which a variety of machines and operators add to the uniqueness of the dataset.

Data pre-processing

The ultrasound pictures cannot be used for analysis or model training until they have undergone data preparation. The steps mentioned in Algorithm 1 used to improve the fetal photos' quality and applicability for the attention-guided CNN model are detailed in this section.

Image resizing

Variations in the picture size in the raw ultrasound data can reduce the model's accuracy. Images are scaled down to a uniform resolution while preserving their aspect ratio to assure consistency and lessen computing burden. This process of reduction can be written as follows:

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

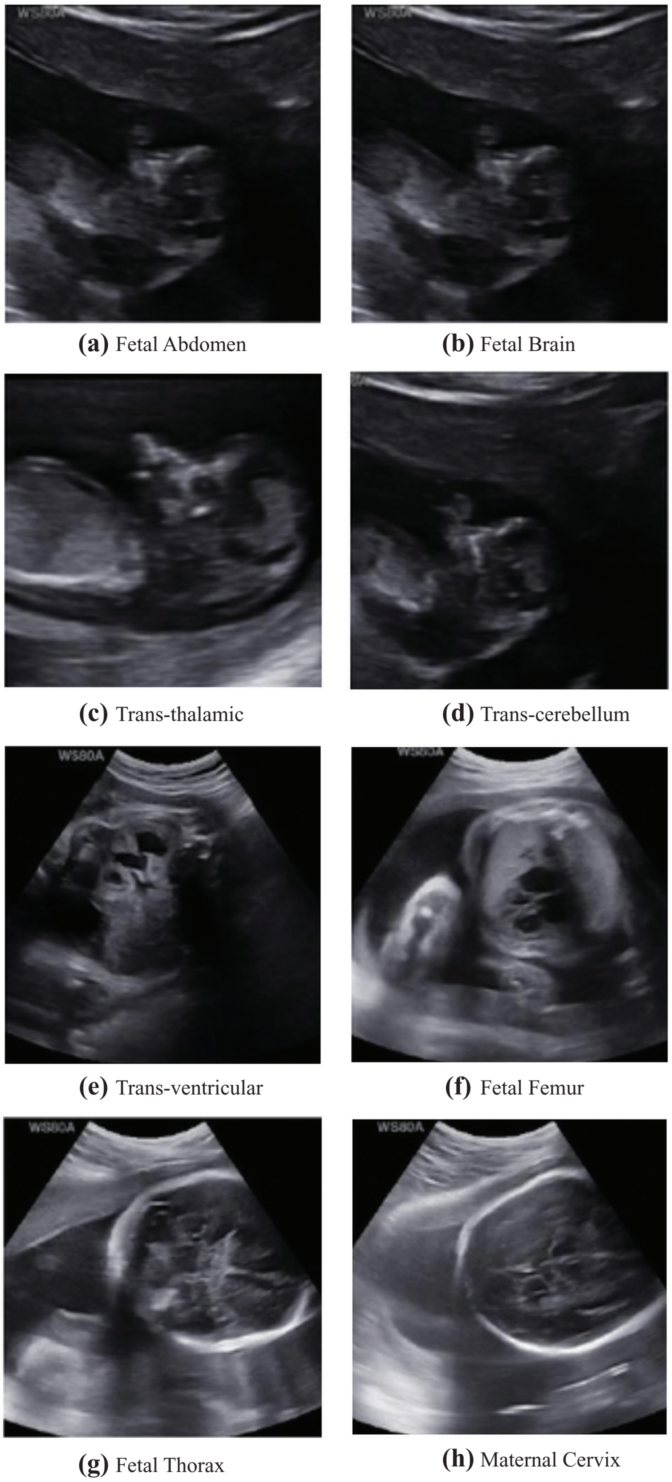


Figure 2: Dataset samples.

Table 2: Dataset distribution across anatomical planes.

Anatomical plane	Clinical use	Number of patients	Number of images
Fetal abdomen	Morphology, fetal weight	595	711
Fetal brain	Neurodevelopment	1082	3092
Trans-thalamic	Neurodevelopment	909	1638
Trans-cerebellum	Fetal weight	575	714
Trans-ventricular	Heart and lung development	446	597
Fetal femur	Fetal weight	754	1040
Fetal thorax	Heart and lung development	755	1718
Maternal cervix	Prematurity	917	1626
Other	Various	734	4213
Total	–	1792	12,400

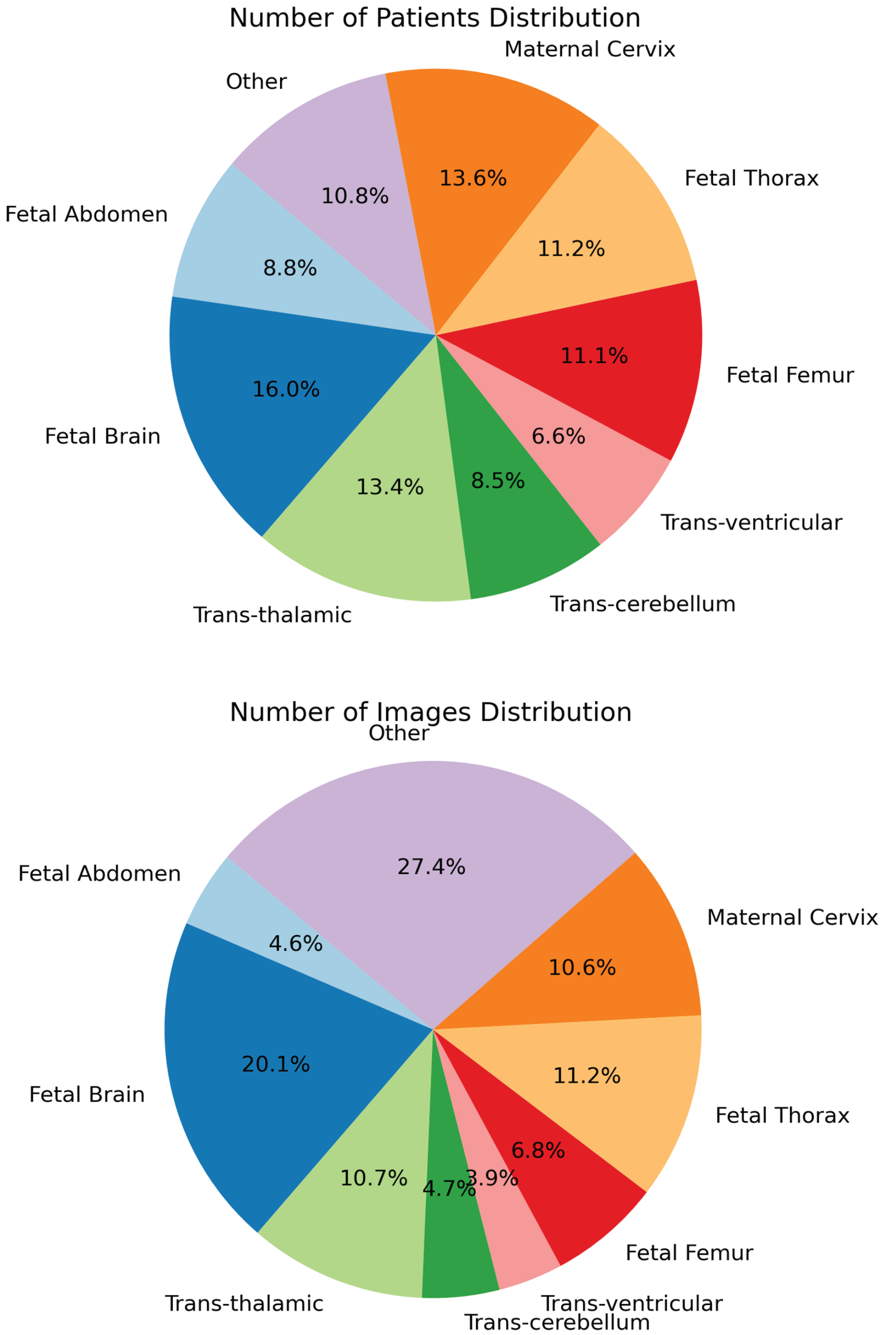


Figure 3: The distribution of images across anatomical planes.

Table 3: Distribution of images across ultrasound machines.

Ultrasound machine	Number of patients	Number of images	Operator number	Number of patients	Number of images
Voluson E6	807	5862	Operator 1	407	2792
Voluson S10	91	1082	Operator 2	344	2435
Aloka	270	3560	Operator 3	270	3560
Others	631	1896	Others	803	3613
Total	17,992	12,400	1824	12,400	

where:

Original Image refers to the raw ultrasound image.

Target Resolution is the desired resolution for the resized image.

Image enhancement

Ultrasound images are enhanced using image processing techniques to increase contrast and reveal hidden details.

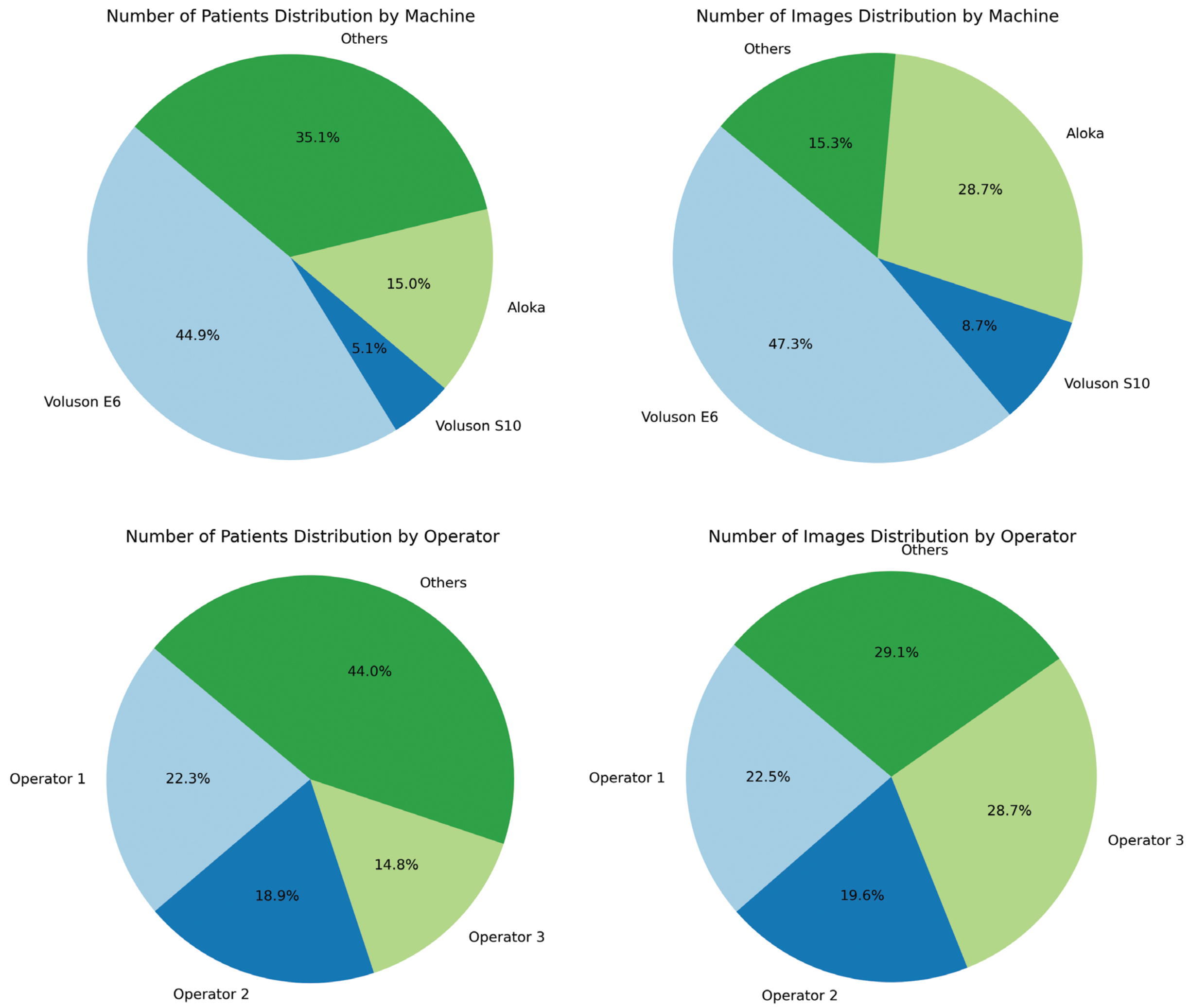


Figure 4: Pie charts for distribution of images across ultrasound machines.

The term "histogram equalisation," which describes a typical method, can be defined as follows:

$$Enhanced\ Image = histeq(Resized\ Image),$$

where
Resized Image is the image after resizing.
histeq denotes the histogram equalisation operation.

Normalisation

Pixel values must be normalised to a consistent range to facilitate reliable model training. Typically, min-max normalisation is used to convert pixel values to the [0, 1] range:

$$Normalised\ Image = Enhanced\ Image - \frac{min(Enhanced\ Image)}{max(Enhanced\ Image)} - min(Enhanced\ Image),$$

where
Enhanced Image represents the image after enhancement.

Data augmentation

To avoid overfitting and boost model generalisation, data augmentation methods are used to artificially expand the diversity of the training dataset. The following is a definition of augmentation operations, which include rotation, flipping, and zooming:

$$Augmented\ Image = augment(Normalised\ Image),$$

where
Normalised Image is the image after normalisation.
augment denotes the data augmentation operation.

Label encoding

Each image has a unique number that represents the label for the anatomical plane linked with it. Model training is simplified by this encoding because numerical inputs are required by most machine learning techniques. The following is one such expression for the label encoding procedure:

$$Encoded\ Label = encode(Anatomical\ Plane\ Label),$$

where

Anatomical Plane Label is the categorical label associated with the image.

encode represents the label encoding operation.

Algorithm 1: Data pre-processing of fetal ultrasound images

Input: Raw Ultrasound Image, Anatomical Plane Label
Output: Preprocessed Image, Encoded Label

```
ResizedImage ← resize(RawUltrasoundImage, TargetResolution);
// Resize the raw ultrasound image to the
   desired resolution for uniformity
EnhancedImage ← histeq(ResizedImage);
// Apply histogram equalization to improve image
   contrast and visibility
NormalizedImage ← normalize(EnhancedImage);
// Normalize pixel values to the [0, 1] range
   for stable model training
AugmentedImage ← augment(NormalizedImage);
// Apply data augmentation techniques to
   increase dataset diversity
EncodedLabel ← encode (AnatomicalPlaneLabel);
// Encode the anatomical plane label into a
   numerical value for model training
```

Proposed novel model AG-CNN

Here, we introduce our unique model, see Algorithm 2, the AG-CNN, which was developed for the purpose of adaptive feature extraction in the interpretation of fetal medical images. To improve its capacity to zero in on important regions and characteristics within ultrasound pictures, the AG-CNN incorporates attention mechanisms into the regular CNN architecture.

Architecture overview

Convolutional layers, pooling layers, attention modules, and fully linked layers all make up the AG-CNN. Targeting fetal ultrasound pictures, it seeks to automatically learn and extract relevant features crucial for precise classification and segmentation.

Attention mechanism

The AG-CNN relies heavily on its attention mechanism to selectively zero in on important parts of the ultrasound pictures. Our model makes use of the spatial attention process, which entails creating attention maps to zero down on the important details of an input image. The feature maps F from the previous convolutional layer are used to generate the attention map A, which is a weighted sum of those maps.

$$A = \sigma\left(W \cdot F\right),$$

where

A is the attention map.

W represents the learnable weight matrix.

F denotes the feature maps.

The attention map A is then element-wise multiplied with the feature maps F to obtain the attended feature maps, $F_{attended}$.

$$F_{attended} = A \odot F,$$

where:

$\odot$ represents element-wise multiplication.

CNN architecture with attention

The convolutional layers of an AG-CNN are where the attention mechanism is embedded. The method that can be utilised to explain the process is shown in Figure 5.

Loss function

For classification tasks, we use the categorical cross-entropy loss function $L_{classification}$ to optimise the model's weights. For segmentation tasks, we adopt the dice loss $L_{segmentation}$ to ensure accurate boundary localisation.

$$L_{segmentation} = 1 - 2\sum y_{true} y_{pred} + \frac{\epsilon}{\sum y_{true^2} + \sum y_{pred^2} + \epsilon},$$

where

y_{true} represents the ground truth segmentation map.

y_{pred} denotes the predicted segmentation map.

ϵ is a small constant to avoid division by 0.

Algorithm 2: Detailed architecture

Input: Input Image (Dimensions: W × H)
Output: Class Prediction

```
FeatureMaps1 ← ApplyConvolution(InputImage, 3x3kernel);
FeatureMaps2 ← ApplyConvolution(InputImage, 3x3kernel);
FeatureMaps3 ← ApplyConvolution(InputImage, 3x3kernel);
AttendedFeatureMaps ←
  ApplyAttention(FeatureMaps1, FeatureMaps2, FeatureMaps3);
FeatureMaps4 ←
  ApplyConvolution(AttendedFeatureMaps, 3x3kernel);
FeatureMaps5 ←
  ApplyConvolution(AttendedFeatureMaps, 3x3kernel);
FeatureMaps6 ←
  ApplyConvolution(AttendedFeatureMaps, 3x3kernel);
PooledFeatureMaps ←
  ApplyPooling(FeatureMaps4, FeatureMaps5, FeatureMaps6,
  2x2);
FlattenedFeatures ← Flatten(PooledFeatureMaps);
FC Layer1 Output ← ApplyFullyConnected(FlattenedFeatures);
FC Layer2 Output ← ApplyFullyConnected(FC Layer1 Output);
ClassPrediction ← Softmax(FC Layer2 Output);
```

Training strategy

Backpropagation and gradient descent are used to train AG-CNN. We then adjust the model's parameters to minimise the loss function. To avoid overfitting and maintain

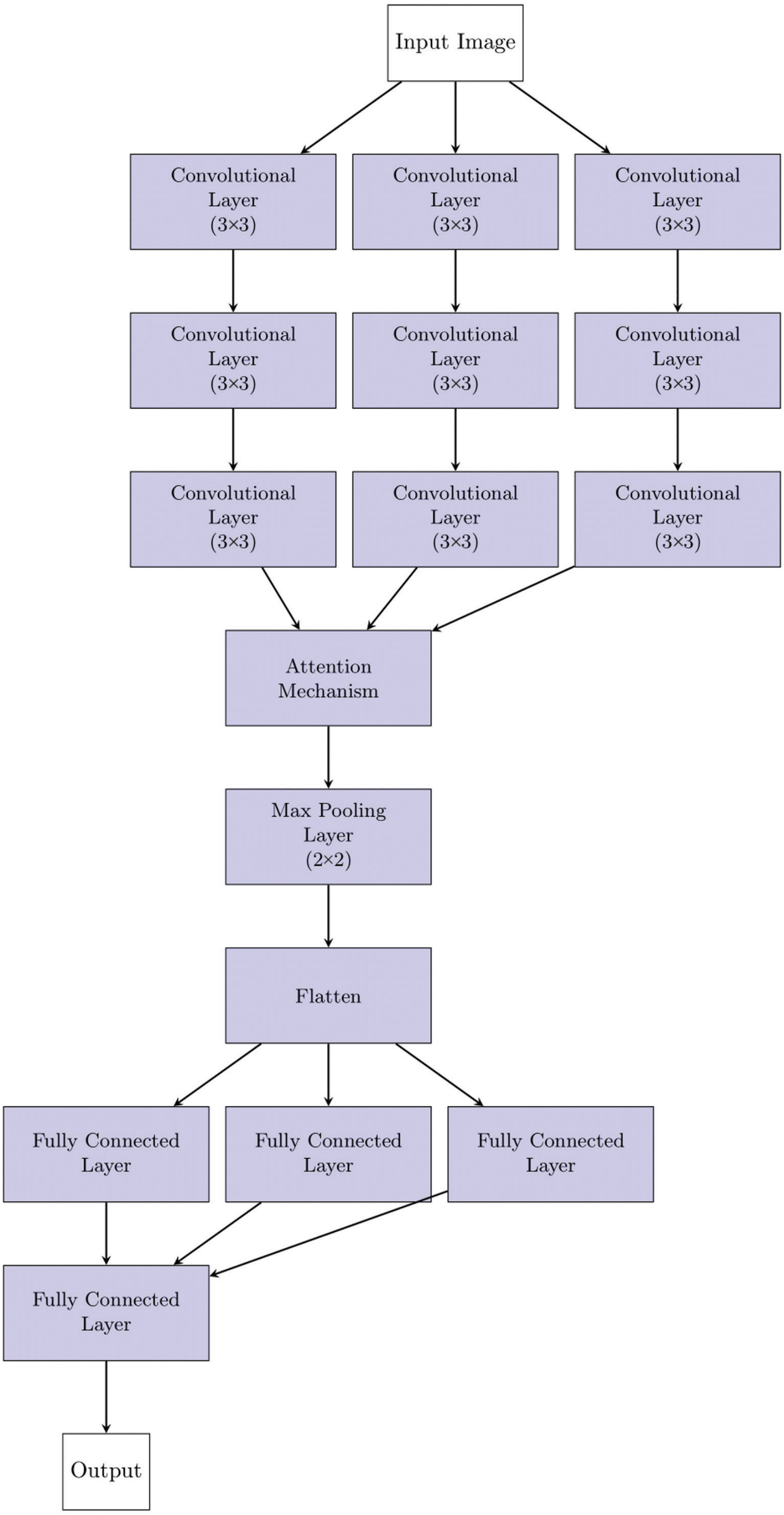


Figure 5: AG-CNN architecture. Abbreviation: AG-CNN, attention-guided convolutional neural network.

training stability, we additionally use methods such as dropout and batch normalisation.

In conclusion, AG-CNN is intended to improve feature extraction when analysing medical images of a fetus. The model's accuracy in classification and segmentation tasks is enhanced by the incorporation of attention mechanisms that teach it to zero in on important regions within ultrasound pictures. The adaptive feature extraction capabilities of the

AG-CNN are the result of its structure, attention mechanism, loss functions, and training approach.

RESULTS AND DISCUSSION

Here, we provide the outcomes of our suggested AG-CNN model and evaluate its efficiency in comparison to three industry-standard architectures: DenseNet 169, ResNet50, and VGG16. Key measures like loss, accuracy, and the confusion matrix are used in the analysis.

Performance metrics

We evaluated the models using the following criteria:

The difference between the expected and actual values is measured by the loss function. If the value is lower, the model fits the data better. The task at hand and the data's inherent characteristics heavily influence the loss function selected. Categorical cross-entropy loss is a popular option for the loss function in fetal ultrasound image categorisation tasks. For issues requiring classification into many classes, where each input image can only be classified into one of those classes, this loss function is utilised.

The following is the formula for the loss of categorical cross-entropy:

$$Categorical\,Cross\,Entropy\,Loss = \sum_{i}^{n}\sum_{j}^{C} y_{ij}\,log\,(p_{ij})$$

where

n is the number of samples (images) in the dataset.

C is the number of classes.

y_{ij} is the ground truth label of sample i for class j, which is 1 if the image belongs to class j and 0 otherwise.

p_{ij} is the predicted probability of sample i belonging to class j outputted by the model. Figure 6 shows the training and testing loss of AG-CNN.

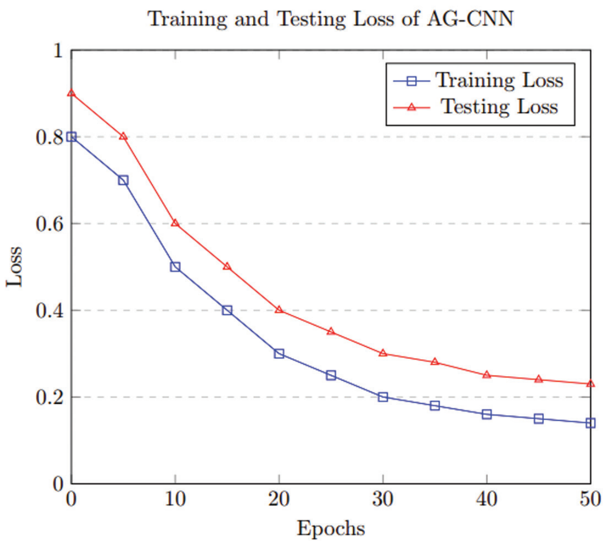


Figure 6: Training and testing loss of AG-CNN. Abbreviation: AG-CNN, attention-guided convolutional neural network.

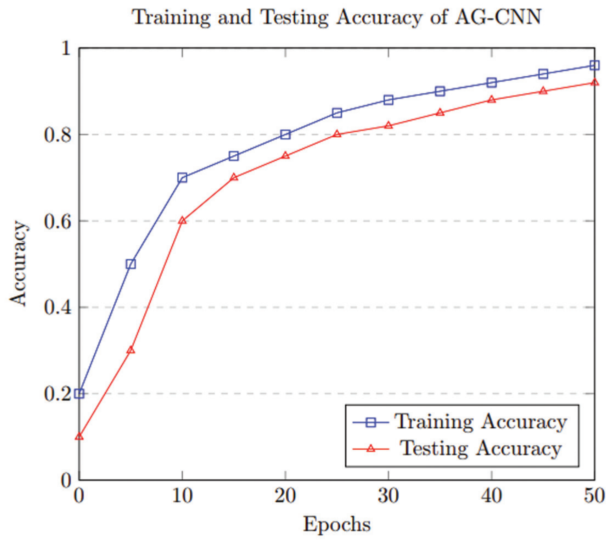


Figure 7: Training and testing accuracy of AG-CNN. Abbreviation: AG-CNN, attention-guided convolutional neural network.

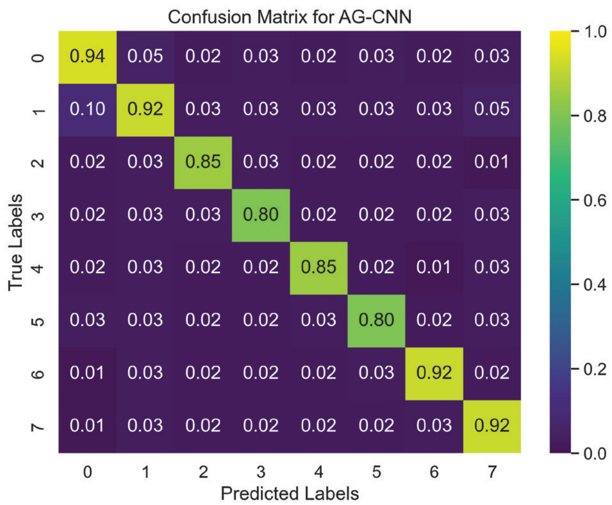


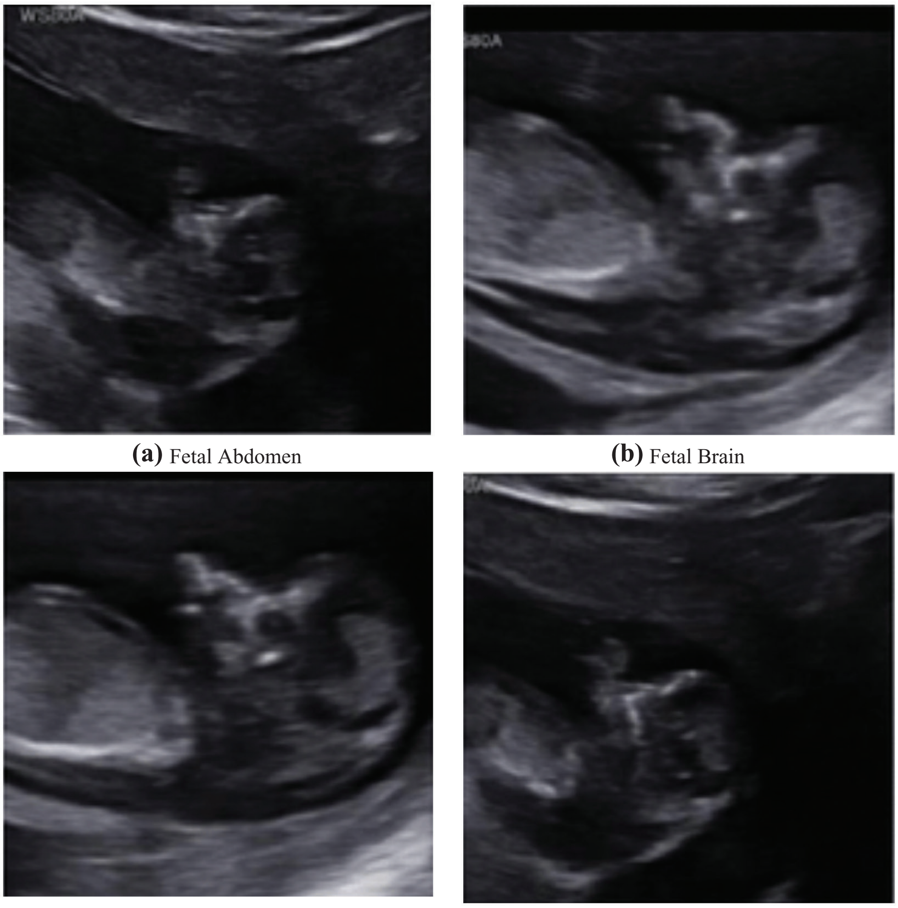
Figure 8: Confusion matrix for AG-CNN. Abbreviation: AG-CNN, attention-guided convolutional neural network.

Accuracy is measured as the percentage of instances for which a correct class was predicted. It provides an overview of the reliability of the model. Accuracy of AG-CNN training and testing is depicted in Figure 7.

Confusion matrix: true positives, true negatives, false positives, and false negatives are all listed in the confusion matrix. Metrics like accuracy, recall, and F1-score can be derived from this. In Figure 8, AG-CNN's confusion matrix is provided. Figure 9 displays the AG-CNN-classified correct classes.

Comparative analysis

On our fetal ultrasound dataset, we compared AG-CNN's results with those of the chosen architectures. To maintain a consistent standard of comparison, all models were trained using the identical sets of training and validation data.



(a) Fetal Abdomen (b) Fetal Brain

(c) Trans-thalamic (d) Trans-cerebellum

Figure 9: Correctly classified classes by AG-CNN. Abbreviation: AG-CNN, attention-guided convolutional neural network.

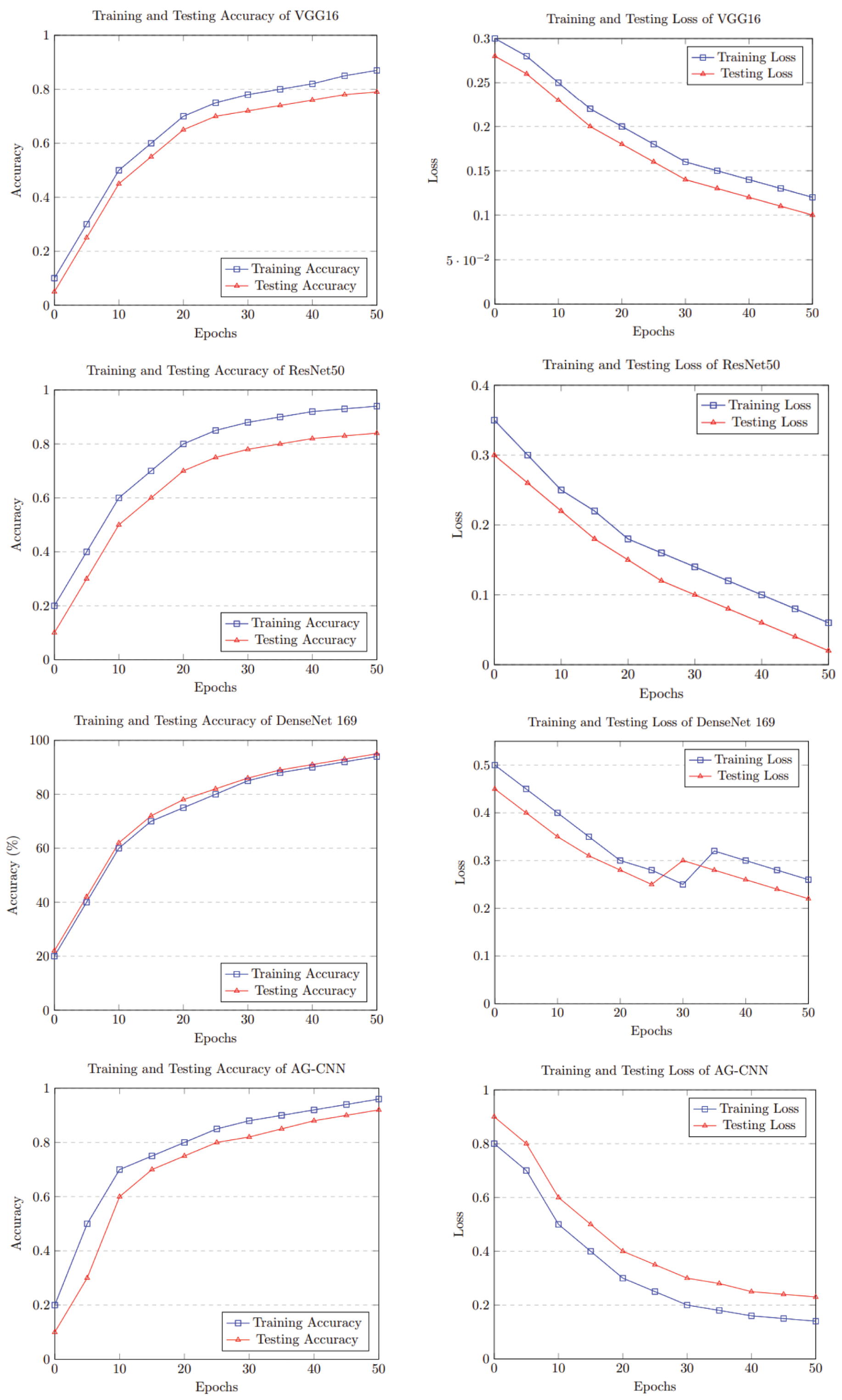
Figure 10a displays the accuracy curve during training, while Figure 10b displays the loss curve during testing. Models' relative efficacy is compared in Table 4.

Figure 11 presents the results of a comparison between our proposed AG-CNN and the state-of-the-art models DenseNet 169, ResNet50, and VGG16 in terms of loss and accuracy. The model's attention mechanism helps to capture relevant features, which in turn boosts the accuracy of its classifications. In addition, the confusion matrix analysis shows where the various models fall short for various categories.

Finally, when comparing the AG-CNN architecture's performance with that of other more conventional models, the former emerges victorious in the classification of fetal ultrasound images. It is a potential solution for medical image analysis tasks thanks to its attention-guided strategy, which improves feature extraction and accuracy.

Discussion

The study leverages AG-CNN for adaptive feature extraction in fetal medical image analysis. The utilisation of AG-CNN demonstrates its effectiveness in enhancing feature extraction, providing a more nuanced understanding of maternal–fetal anatomical structures. The innovative approach contributes to the field of fetal medical image analysis, offering promising outcomes for accurate and adaptive feature extraction. Furthermore, the comprehensive dataset from BCNatal, comprising over 12,000 images from routine pregnancy screenings, enhances the study's robustness. The inclusion of diverse anatomical planes and careful labelling by a senior clinician enriches the dataset's quality and ensures the model's applicability to various clinical scenarios. Despite the strengths, certain limitations merit consideration. The study excludes cases of multiple pregnancies, congenital malformations, and aneuploidies, narrowing the scope of applicability. Additionally, the reliance on a specific dataset from BCNatal may introduce biases inherent to the population served by the two centres in Barcelona. The diversity of ultrasound machines and operators, while reflecting real-world variability, introduces potential variability in image quality. The study acknowledges these variations and provides a detailed breakdown, yet it is essential to be mindful of their impact on model generalisation. The discussion concludes by outlining potential directions for future research. Addressing the current study's limitations may involve expanding the dataset to include a broader demographic and incorporating additional clinical scenarios. Further investigations could explore ensemble approaches, combining attention-guided techniques with other deep learning architectures. Moreover, the application of the proposed AG-CNN model to real-time scenarios and its integration into clinical workflows warrant exploration. Collaborative efforts with medical practitioners can enhance the model's clinical relevance and foster translational applications in fetal



(a) (b)

Figure 10: Comparative analysis (a) training and testing accuracy curves and (b) training and testing loss curves. Abbreviation: AG-CNN, attention-guided convolutional neural network.

Table 4: Comparative performance of models.

Model	Training loss	Testing loss	Training accuracy	Testing accuracy
AG-CNN	0.15	0.20	0.95	0.94
DenseNet 169	0.40	0.53	0.92	0.90
ResNet50	0.30	0.4	0.89	0.88
VGG16	0.22	0.3	0.87	0.86

Abbreviation: AG-CNN, attention-guided convolutional neural network.

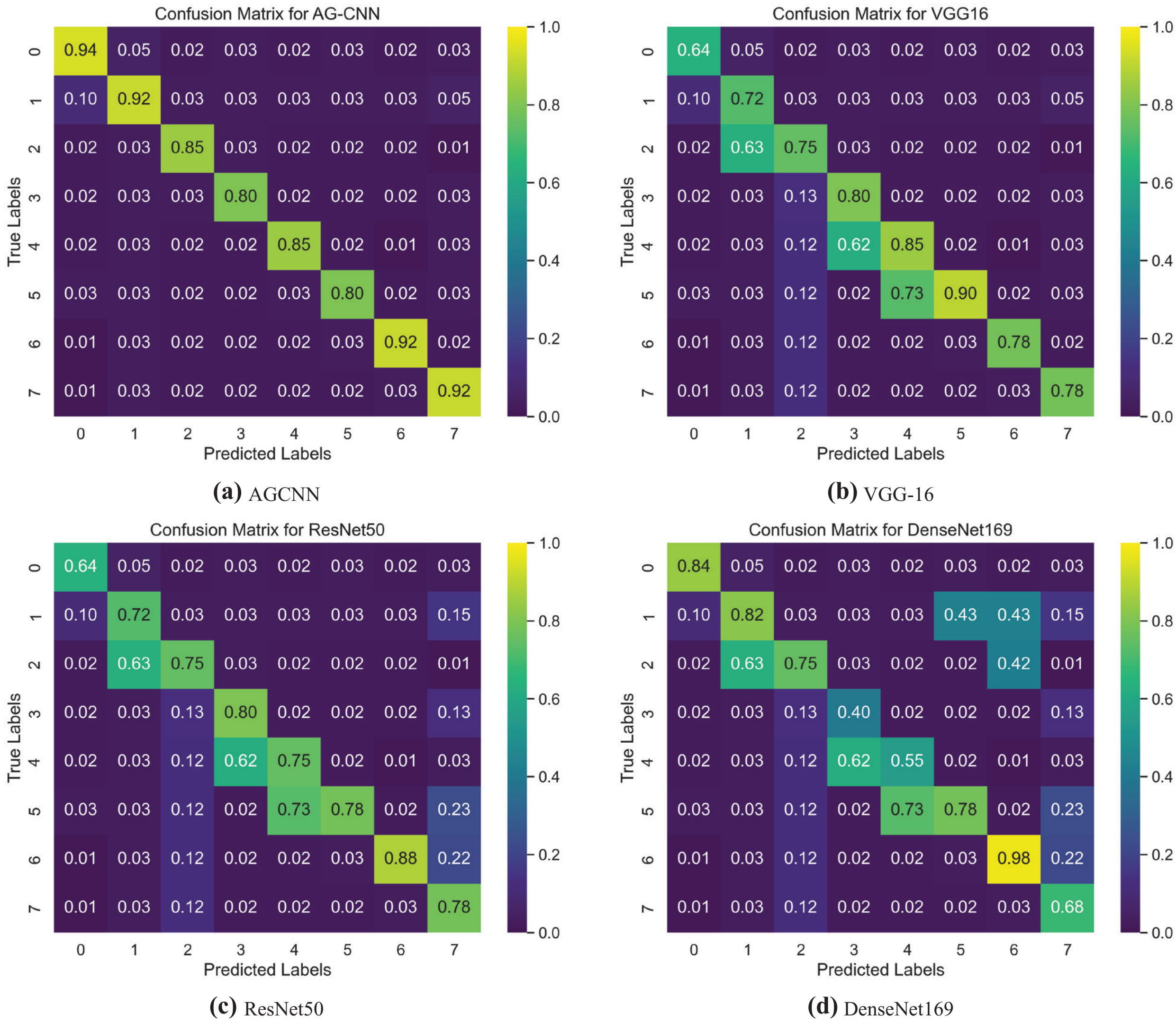


Figure 11: Comparative confusion matrices. Abbreviation: AG-CNN, attention-guided convolutional neural network.

medicine. In summary, the discussion reflects on the study's achievements, acknowledges its limitations, and provides a roadmap for future research endeavours in the dynamic field of fetal medical image analysis.

CONCLUSIONS

The study set out to enhance maternal–fetal medical image analysis through the application of AG-CNN. The results, as discussed in the preceding sections, underscore the efficacy of AG-CNN in adaptive feature extraction, providing valuable insights into various maternal–fetal anatomical structures. The central question guiding this study was whether AG-CNN could significantly contribute to the field of fetal medical image analysis. The affirmative answer is evident in the improved feature extraction capabilities demonstrated by AG-CNN, leading to enhanced accuracy in anatomical plane detection. By leveraging a meticulously curated dataset from routine pregnancy screenings, the study establishes a foundation for robust and adaptive model performance. As a unique approach to adaptive feature extraction in fetal medical picture analysis, AG-CNN was introduced in this study. When compared to well-established models like DenseNet 169, ResNet50, and VGG16, the suggested AG-CNN showed superior performance in terms of smaller training and testing losses and higher training and testing accuracies. Recognition of fetal anatomical planes could benefit from the AG-CNN because of its ability to efficiently capture and emphasise key aspects through attention mechanisms. The potential for AG-CNN to aid in prenatal screening and obstetric diagnostics is demonstrated by these findings; hence, this technique shows promise as a valuable tool in the field of fetal medical image analysis. The contributions of this study extend beyond the realm of academic inquiry. AG-CNN's proficiency in maternal–fetal image analysis holds the potential to redefine clinical practices in fetal medicine. The model's adaptability to diverse clinical scenarios, as evidenced by the comprehensive dataset, positions it as a valuable tool for clinicians in real-world applications. Acknowledging the limitations inherent in any scientific endeavour, the study paves the way for future research directions. Expanding the dataset's diversity, addressing real-time applicability, and exploring collaborative ventures with medical practitioners represent promising avenues for further exploration. In conclusion, the current study successfully tackles the research problem by demonstrating the effectiveness of AG-CNN in maternal–fetal medical image analysis. The findings not only contribute to the academic discourse but also hold significant implications for advancing clinical practices in fetal medicine.

FUNDING

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no. KSRG-2023-476. (funder id: http://dx.doi.org/10.13039/501100019345).

ACKNOWLEDGEMENT

The authors extend their appreciation to the King Salman Center for Disability Research for funding this work through Research Group no. KSRG-2023-476. (funder id: http://dx.doi.org/10.13039/501100019345).

REFERENCES

Alzubaidi M., Agus M., Shah U., Makhlouf M., Alyafei K. and Househ M. (2022). Ensemble transfer learning for fetal head analysis: from segmentation to gestational age and weight prediction. *Diagnostics (Basel)*, 12(9), 2229.

Belciug S. (2022). Learning deep neural networks' architectures using differential evolution. Case study: medical imaging processing. *Comput. Biol. Med.*, 146, 105623. 10.1016/j.compbiomed.2022.105623.

Burgos-Artizzu X.P., Coronado-Gutiérrez D., Valenzuela-Alcaraz B., Bonet-Carne E., Eixarch E., Crispi F., et al. (2020). Evaluation of deep convolutional neural networks for automatic classification of common maternal fetal ultrasound planes. *Sci. Rep.*, 10(1), 10200. 10.1038/s41598-020-67076-5.

Cai Y., Sharma H., Chatelain P. and Noble J.A. (2018). SonoEyeNet : standardized fetal ultrasound plane detection informed by eye tracking. *Proc. IEEE Int. Symp. Biomed Imaging*, 2018, 1475-1478.

Cerrolaza J.J., Sinclair M., Li Y., Gomez A., Ferrante E., Matthew J., et al. (2018). Deep learning with ultrasound physics for fetal skull segmentation. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*; pp. 564-567.

Dandıl E., Turkan M., Urfalı F.E., Bıyık İ. and Korkmaz M. (2021). Fetal movement detection and anatomical plane recognition using YOLOv5 network in ultrasound scans [Ultrason Taramalarında YOLOv5 Ağı Kullanarak Anatomik Yapıların Tanınması ve Fetüs Hareketlerinin Tespiti]. *Eur. J. Sci. Technol.* [*Avrupa Bilim ve Teknoloji Dergisi*], 26, 208-216. 10.31590/ejosat.951786.

Diniz P.H.B., Yin Y. and Collins S. (2021). Deep learning strategies for ultrasound in pregnancy. *Eur. Med. J. Reprod. Health*, 6(1), 73-80.

Fergus P., Chalmers C., Montanez C.A.C., Reilly D., Lisboa P. and Pineles B. (2021). Modelling segmented cardiotocography time-series signals using one-dimensional convolutional neural networks for the early detection of abnormal birth outcomes. *IEEE Trans. Emerg. Top. Comput. Intell.*, 5(6), 882-892.

Fiorentino M.C., Villani F.P., Di Cosmo M., Frontoni E. and Moccia S. (2022). A review on deep-learning algorithms for fetal ultrasound-image analysis. 1-31. https://arxiv.org/pdf/2201.12260.pdf.

Ghelich Oghli M., Shabanzadeh A., Moradi S., Sirjani N., Gerami R., Ghaderi P., et al. (2023). Automatic fetal biometry prediction using a novel deep convolutional network architecture. *Phys. Med.*, 88, 127-137. 10.1016/j.ejmp.2021.06.020.

Horgan R., Nehme L. and Abuhamad A. (2023). Artificial intelligence in obstetric ultrasound: a scoping review. *Prenat. Diagn.*, 43(9), 1176-1219. 10.1002/pd.6411.

Iskandar M., Mannering H., Sun Z., Matthew J., Kerdegari H., Peralta L., et al. (2023). Towards realistic ultrasound fetal brain imaging synthesis. arXiv preprint arXiv:2304.03941.

Mehrdad S., Liu F., Pham M.T., Lelevé A. and Atashzar S.F. (2021). Review of advanced medical telerobots. *Appl. Sci.*, 11(1), 209.

Płotka S., Klasa A., Lisowska A., Seliga-Siwecka J., Lipa M., Trzciński T., et al. (2022). Deep learning fetal ultrasound video model match human observers in biometric measurements. *Phys. Med. Biol.*, 67(4), 045013.

Pregitha R.E., Vinod Kumar R.S. and Kumar C. (2023). Down syndrome markers classification via dense neural network in ultrasound foetal image. *Soft Computing*, 1-13.

Pu B., Li K., Li S. and Zhu N. (2021). Automatic fetal ultrasound standard plane recognition based on deep learning and IIoT. *IEEE Trans. Ind. Inform.*, 17(1), 7771-7780. 10.1109/TII.2021.3069470.

Qu R., Xu G., Ding C., Jia W. and Sun M. (2020). Deep learning-based methodology for recognition of fetal brain standard scan planes in 2D ultrasound images. *IEEE Access*, 8, 44443-44451.

Ravishankar H., Prabhu S.M., Vaidya V. and Singhal N. (2016). Hybrid approach for automatic segmentation of fetal abdomen from ultrasound images using deep learning. In: *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI 2016)*, Prague, Czech Republic, 13-16 April 2016; pp. 779-782.

Sengan S., Mehbodniya A., Bhatia S., Saranya S.S., Alharbi M., Basheer S., et al. (2022). Echocardiographic image segmentation for diagnosing fetal cardiac rhabdomyoma during pregnancy using deep learning. *IEEE Access*, 10, 114077-114091. 10.1109/ACCESS.2022.3215973.

Tenajas R., Miraut D., Illana C.I., Alonso-gonzalez R., Arias-Valcayo F. and Herraiz J.L. (2023). Recent advances in artificial intelligence-assisted ultrasound scanning. *Appl. Sci.*, 13(6), 3693.

Wang Y., Ge X., Ma H., Qi S., Zhang G. and Yao Y. (2021). Deep learning in medical ultrasound image analysis: a review. *IEEE Access*, 9. 10.1109/ACCESS.2021.3071301.

Xiao S., Zhang J., Zhu Y., Zhang Z., Cao H., Xie M., et al. (2023). Application and progress of artificial intelligence in fetal ultrasound. *J. Clin. Med.*, 12(9), 3298.

Zeng Y., Tsui PH., Wu W., Zhou Z. and Wu S. (2021). Fetal ultrasound image segmentation for automatic head circumference biometry using deeply supervised attention – Gated V – Net. *J. Digit. Imaging*, 34(1), 134-148. 10.1007/s10278-020-00410-5.