



Fetal Movement Detection and Anatomical Plane Recognition using YOLOv5 Network in Ultrasound Scans

Emre Dandıl^{1*}, Musa Turkan², Furkan Ertürk Urfalı³, İsmail Bıyık⁴, Mehmet Korkmaz⁵

¹ Bilecik Seyh Edebali University, Faculty of Engineering, Department of Computer Engineering, Bilecik, Turkey (ORCID: 0000-0001-6559-1399)

² Bilecik Seyh Edebali University, Institute of Graduate, Department of Computer and Electronics Engineering, Bilecik, Turkey (ORCID: 0000-0002-4370-7474)

³ Kütahya Health Sciences University, Evliya Çelebi Training and Research Hospital, Department of Radiology, Kütahya, Turkey (ORCID: 0000-0002-4875-7761)

⁴ Kütahya Health Sciences University, Evliya Çelebi Training and Research Hospital, Department of Obstetrics and Gynecology, Kütahya, Turkey (ORCID: 0000-0001-6111-9302)

⁵ Kütahya Health Sciences University, Evliya Çelebi Training and Research Hospital, Department of Radiology, Kütahya, Turkey (ORCID: 0000-0001-6234-9484)

(International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA) 2021 – 11-13 June 2021)

(DOI: 10.31590/ejosat.951786)

ATIF/REFERENCE: Dandıl, E., Turkan, M., Urfalı, F. E., Bıyık, İ. & Korkmaz, M. (2021). Fetal Movement Detection and Anatomical Plane Recognition using YOLOv5 Network in Ultrasound Scans. *Avrupa Bilim ve Teknoloji Dergisi*, (26), 208-216.

Abstract

Analyzing medical images and videos with computer-aided algorithms provides important benefits in the diagnosis and treatment of diseases. Especially in recent years, the increasing developments in deep learning algorithms have provided continuous improvement in subjects such as speed, performance and hardware need in the processing of medical data. Examination of medical data, which may require advanced expertise, using deep learning algorithms has begun to be widely used as a secondary tool in the decision-making process of physicians. Tracking the movements of the fetus and recognizing its planes in ultrasound (US) videos is an important parameter in evaluating the health of the baby. In this study, a YOLOv5 deep learning network based method is proposed to identify fetal anatomical planes from fetal ultrasound and to detect their movements. First of all, a dataset of videos containing 16-20 weeks of fetal movements is created in the study. In the next step, the fetal head, arm, heart and body are identified and tracking using the deep-SORT algorithm on the labeled data. In the experimental studies conducted on ultrasound videos within the scope of the study, using the YOLOv5 algorithm, head, body, heart and arm are recognized with 95.04%, 94.42%, 88.31% and 83.23% F1-score, respectively. In addition, ultrasonic video movements of the head, heart and body of the fetus are followed and the trajectories and patterns of the movements are extracted. Thus, the detection of fetal movements from the movement patterns transformed into a two-dimensional plane is achieved.

Keywords: Fetus, Ultrasound video, Deep learning, Anatomical plane recognition, Fetal movement detection, YOLOv5, Deep-SORT algorithm, Object tracking.

Ultrason Taramalarında YOLOv5 Ağı Kullanarak Anatomik Yapıların Tanınması ve Fetüs Hareketlerinin Tespiti

Öz

Tıbbi görüntülerin ve videoların bilgisayar destekli algoritmalar ile analiz edilmesi, hastalıkların teşhis ve tedavisinde önemli faydalar sağlamaktadır. Özellikle son yıllarda, derin öğrenme algoritmalarındaki artan gelişmeler, medikal verilerin işlenmesinde hız, performans ve donanım ihtiyacı gibi konularda sürekli iyileşme sağlamıştır. İleri derece uzmanlık gerektirebilen medikal verilerin inceleme işlemlerinin derin öğrenme algoritmalarıyla yapılması, hekimlerin karar verme aşamasında ikincil bir araç olarak yaygın olarak kullanılmaya başlamıştır. Ultrason (US) videolarında fetüsün hareketlerinin izlenmesi ve anatomik yapıların tanınması bebek sağlığının değerlendirilmesinde önemli bir parametredir. Bu çalışmada, fetüs ultrasonundan fetüs anatomik yapıların tanınması ve hareketlerinin tespiti için YOLOv5 derin öğrenme ağı destekli bir yöntem önerilmektedir. Çalışmada öncelikle, 16-20 haftalık fetüs hareketlerini içeren videolardan bir veriseti oluşturulmuştur. Sonraki aşamada, etiketlenen veriler üzerinde deep-SORT algoritması kullanılarak, fetüse ait kafa, kol, kalp ve gövde kimliklendirilmiş ve takip edilmiştir. Çalışma kapsamında ultrason videolarında

* Corresponding Author: Bilecik Seyh Edenali University, Faculty of Engineering, Department of Computer Engineering, Bilecik, Turkey (ORCID: 0000-0001-6559-1399), emre.dandil@bilecik.edu.tr

yürütülen deneysel çalışmalarda, YOLOv5 algoritması kullanılarak, kafa, gövde, kalp ve kol anatomik yapıları, sırasıyla %95.04, %94.42, %88.31 ve %83.23 F1 skoru ile tanınmıştır. Ayrıca, fetüse ait kafa, kalp ve gövdenin ultrasonik video hareketleri takip edilerek, hareketlerin yörüngeleri ve örüntüleri çıkarılmıştır. Böylece, iki boyutlu düzleme dönüştürülen hareket örüntülerinden fetüs hareketlerinin tespiti sağlanmıştır.

Anahtar Kelimeler: Fetüs, Ultrason video, Derin öğrenme, Anatomik yapı tanıma, Fetal hareket tespiti, YOLOv5, Deep-SORT algoritması, Nesne takibi.

1. Introduction

Fetal biometric data can provide strong diagnostic and prognostic indications of fetal development (Fiorentino *et al.*, 2021). In addition to monitoring the development of organs such as the head, body and heart of the fetus, tracking of fetal movements is very important in terms of maternal and infant health. In recent years, abnormal fetal development has led to an increase in prenatal and natal mortality in both developing and industrialized countries (Salomon *et al.*, 2011). Therefore, detection of fetal abnormalities at an early stage stands out in that it is noteworthy.

Some diseases can be diagnosed early by investigating and evaluating fetal scans. 2D ultrasound imaging is primarily used to assess fetal health because of its low cost, comprehensive availability, real-time capability, and no harmful radiation (Baumgartner *et al.*, 2017). Ultrasound imaging is an imaging method frequently used in gestation follow-up, as it is an important parameter that provides information about the health status of the baby, thanks to the monitoring of fetal movements. In addition, fetal development and fetal movements can be observed using real-time ultrasound imaging. Evaluation of fetal videos and images can be performed by specialist physicians. However, identification and monitoring of fetal movements from fetal video and images often involves a difficult and complex process.

Currently, routine ultrasound scans are performed at least once between the 18th and 20th weeks of pregnancy in most countries (Salomon *et al.*, 2011). In these scans, some biometric measurements and evaluations are typically applied for the anatomy of the fetus, such as measuring head circumference, monitoring arm and leg bone lengths, listening to heartbeat. In addition, biometric measurements such as fetal weight and abdominal circumference are also used to diagnose the health status of the fetus (Deepika *et al.*, 2021). Moreover, following the movements of the fetus is an important parameter in a healthy gestation process. Reduced or absent fetal movements in the mother's womb may be a sign of serious fetus risks (You *et al.*, 2017). It has become available in recent years to collect information on how the fetus moves and on quantitative and qualitative movement patterns during pregnancy. Fetal movements can be classified as general movements and movements that may cause anxiety. While general movements indicate a healthy pregnancy process, worrisome movements may be a sign of poor fetal health (Bai, 2016). It has been stated that during the first trimester of pregnancy, the measurements obtained from ultrasound scans are used to predict the risky gestational age (Looney *et al.*, 2017). Two-dimensional images are produced by placing an ultrasound transducer on the womb along the fetal thorax (between the abdomen and neck) and the abdominal axis. Generally, an observation period of 5-30 minutes is considered sufficient. Using ultrasound images, critical anatomical structures can be detected, fetal movements can be tracked, and the gestation can be monitored. Frequency, intensity and duration of fetal movements are related to maternal perception of movements. Fetal movements are also an

important process as they are used to detect an inactive heartbeat condition (Wróbel *et al.*, 2014). Although fetal movements can be observed using real-time ultrasound imaging, some studies have been carried out to determine the pattern of fetal movements, but a general framework has not been defined for the amount of these movements, the ideal number, and when they should be counted (Velazquez & Rayburn, 2002).

Increased fetal movements may not always be a good sign for the pregnancy period. In some cases, increased fetal movements may be a sign of undesirable conditions. In this case, it is difficult to determine which kinds of movements are undesirable. Some of the excessive movement may be a sign of serious fetal disorders. Since it is difficult for women to understand this, intervention may be delayed. Pregnant women feel a complex series of fetal movement patterns throughout pregnancy. Many factors, including maternal obesity, maternal behavior and mood, affect women's perception of fetal movement (Whitehead *et al.*, 2020).

Obtaining fetal biometric measurements through two-dimensional ultrasound scans is crucial for fetal monitoring. However, obtaining standardized head, abdomen, and femoral image planes is difficult due to variable image quality (Ahmed & Noble, 2016). For this reason, many ultrasound image classification studies have been carried out to assist experts in their evaluation (Carneiro *et al.*, 2008; Ishikawa *et al.*, 2019; Kumar *et al.*, 2016; Malathi & Shanthi, 2009; Sinclair *et al.*, 2018; Sobhaninia *et al.*, 2019; Yaqub *et al.*, 2012).

When the studies in the literature are investigated, it is seen that there are studies that suggest examining fetal US scans for different purposes and with different methods and aiming to provide convenience to physicians (Carneiro *et al.*, 2008). In addition, studies have also been proposed to performed biometric measurements in the fetus to determine the organs, the standard plane of the fetus (Lei *et al.*, 2014; Yaqub *et al.*, 2012; Yu *et al.*, 2016). Moreover, it was observed that the researchers mostly used their own datasets in the studies.

There are many studies using deep learning methods for the examination and evaluation of fetal ultrasound images (Gao *et al.*, 2016; Ravishankar *et al.*, 2016; Sinclair *et al.*, 2018; van den Heuvel *et al.*, 2019; Yu *et al.*, 2016). The performance increase that occurs with deep learning depends on the large datasets used for training. Although there are studies that provide high-performance results with developing technology and the application of different deep learning methods, it is still seen that progress needs to be carried on. In addition, although there are publicly available datasets in different fields, the datasets that can be accessed especially in terms of fetuses are limited. The limited open fetal ultrasound dataset is one of the most important constraints in the development of common deep learning applications in this field. Researchers working with small datasets use the transfer learning method that enables cross learning in different datasets to overcome this problem. A pre-trained network is used to extract features directly and fine-tuning is done by fixing the weights in some parts of the network. The use of pre-trained networks using large-scale

natural images increases the performance of networks trained with low resolution ultrasound images (Gao *et al.*, 2016).

Identification of objects in fetus US scans is important for extracting trajectories from video images. Object identification is a different process from object recognition. It requires naming two different objects in the image as different objects. When each object in an image is named with a different identity, the motion trajectories of each object can be recorded. Obtaining the trajectories of fetal movements may allow the recognition of different movement types. In this study, a deep learning-based approach is proposed for the recognition of some anatomical structures (planes) belonging to the fetus such as head, arm, heart from fetal ultrasound and the detection of their movement trajectories. In this study, fetal anatomical planes are identified and tracked using YOLOv5 network and the deep-SORT algorithm in the dataset created from videos containing fetal movements. In this study, it was observed that the motion trajectory information formed different patterns according to the organs. If enough video images with healthy and unhealthy movements are obtained, it may be possible to classify these movements from orbital images. The next sections of the study are organized as follows. In the second section, the ultrasound dataset and the proposed YOLOv5 network are detailed. The research results and discussion reached as a result of the experimental studies are explained in the third section. In the last part, the inferences obtained within the scope of the study and the experimental measurement results are evaluated and the following works in the future are mentioned.

2. Material and Method

2.1. Fetus Ultrasound Dataset

Fetus scans used in this study were obtained from Evliya Çelebi Training and Research Hospital of Kütahya Health Sciences University. Ultrasound scans were collected from pregnant volunteers between 16 and 20 weeks. All pregnant women were selected from second trimester cases, as it allows easy evaluation of fetal movements. Head and body movements, major extremity movements, 4-chamber heart movements, diaphragm and swallowing movements in all fetuses were recorded as video. US evaluations were performed using a US system (Acuson S3000; Siemens Medical Solutions, Mountain View, CA), which includes a convex transducer that enables scanning with 4 MHz frequency and obstetric US preset. Ultrasound scans were acquired at between 20 and 27 fps. A total of 1648 2D ultrasound images, 1300 for training the network and 348 for testing, were acquired and labeled with the consensus of 3 different experts. Frames were obtained from the video images provided for use in experimental studies in order to train the proposed YOLOv5 network. On the videos in this dataset, the trajectories of the movements of the anatomical structures of the fetus such as head, arm and body were extracted. The sample scans from the fetal ultrasound dataset created within the scope of the study are shown in Figure 1.

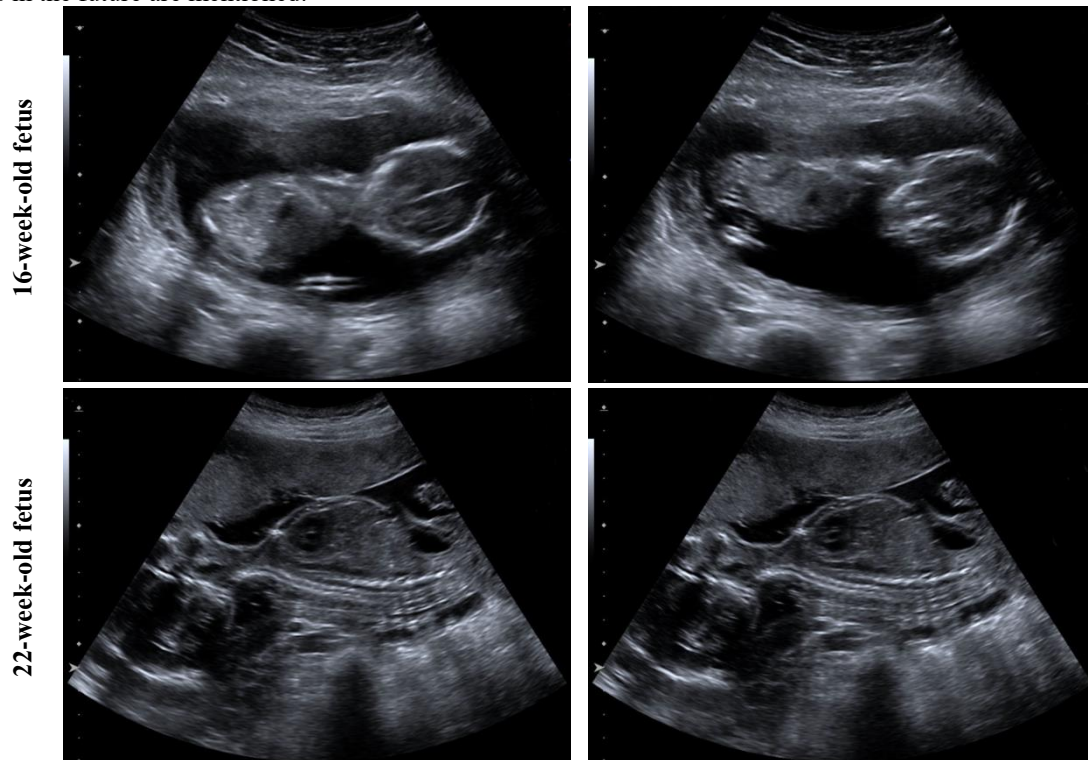


Figure 1. US scans of a 16-week-old fetus and a 22-week-old fetus on the fetus US dataset

2.2. YOLOv5 (you only look once) Deep Learning Network

Deep learning algorithms can be classified in two different ways as single-stage methods (YOLO, SSD, etc.) and two-stage methods (R-CNN, Faster R-CNN etc.). In two-stage algorithms, classification and location determination processes are performed separately, while in single-stage algorithms, both location information and classification can be done by

processing an image. Although higher detection accuracy can be achieved with two-stage algorithms, they are generally slower (Soviany & Ionescu, 2018). A proper balance has to be provided between accuracy and speed when using algorithms. For real-time applications, single-stage methods are preferred for real-time applications, as speed is a prominent evaluation factor.

YOLO (you only look once) is a deep learning algorithm in scope of the single-stage algorithms. The YOLO algorithm is the

deep learning model proposed by (Redmon *et al.*, 2016). The YOLO algorithm has made a difference with its ability to detect objects much faster than previous algorithms. The algorithm can both classify objects and detect the positions of objects by passing the image through the convolutional neural network only once. It also gives fast results because it passes the image through the network only once. This feature makes the YOLO algorithm stand out in real-time object detection. It is seen that the YOLO model works faster than other object recognition methods. The YOLOv5 model was published by Glenn Jocher in 2020 based on PyTorch. YOLOv5 promises higher FPS speed than previous YOLO versions. Also, this network generates lower model files and the training time is shorter. Model weight file size may vary in different models of YOLOv5 (Jocher, 2020). Developed with the effort of 58 open source contributors, YOLOv5 offers some different configurations for object recognition, unlike previous versions (Malta *et al.*, 2021).

YOLOv5, a typical one-stage object recognition network, consists of three modular components: backbone, neck and head (Liu *et al.*, 2021). In the YOLOv5 model, the backbone is mainly used to extract important features from the input image. DarkNet framework-based Cross Stage Partial Networks (CSPNet) is used as the backbone in the YOLOv5 algorithm (Wang *et al.*, 2020). The neck component of the YOLOv5 model is used to generate the feature pyramid network (FPN). Feature pyramids help to generalize well in object scaling. It also helps to recognize the same object at different sizes and resolutions. There are many attributes pyramids of different types. In the YOLOv5 model, the path aggregation network (PANet) is generally used as neck to obtain the pyramid features. In the head, the last layer of YOLOv5, object recognition is performed. In this layer, object class probabilities and bounding boxes are generated. In YOLOv5, model head operations are performed in the same way as in YOLOv4. The YOLOv5 architecture contains four different models such as YOLOv5s, YOLOv5m, YOLOv5l and YOLOv5x. There are fundamental differences between these model versions, such as the number of feature extraction modules and convolution kernels in the structure of these networks (Yan *et al.*, 2021). In this study, YOLOv5s, the smallest version of the YOLOv5 model, was used to recognize some fetal anatomical structures such as head, arm, and heart from fetal ultrasound.

2.3. Deep-SORT Algorithm

Identification and tracking of objects from video frames is a different process than object recognition. It requires an identification of an object moving on the screen which objects in the previous frame. In processes such as counting vehicles in traffic and observing people in crowded environments, each person has to be tracked with a separate identity. In this regard, SORT (simple online and real-time tracking) approach is one of the most successful algorithms that provide identification of objects for tracking. This algorithm can determine which objects in each frame are which objects in the previous frame, using the Kalman filter. It can quickly identify objects using Kalman filter and Hungarian algorithm. It works faster than other tracking techniques. This feature makes it preferred for real-time object identification and tracking (Bewley *et al.*, 2016). Deep-SORT algorithm is an extension of the SORT algorithm used for object

tracking with the addition of a deep appearance descriptor. In the Deep-SORT algorithm, objects can be tracked in longer periods (Wojke *et al.*, 2017).

In this study, Deep-SORT algorithm was used to track movements by identifying organs and anatomical structures in fetal ultrasound. In the study, first of all, while the movement trajectories of the fetal organs are extracted, the trajectory information of the left and right arms have to be recorded separately and identified so that they can be examined. In order to achieve this, the deep-SORT algorithm was used. In this study, objects were detected using the YOLOv5 algorithm, and movements were tracked with the deep-SORT algorithm. The position of the moving object on the screen is recorded in each frame. Orbital information was recorded as a point (x, y) on a two-dimensional plane. By adding the time dimension to the orbit information, it can be evaluated as a three-dimensional data. In the study, while examining video sequences in many experiments, it was ensured that the movements were defined by taking into account the time dimension. The trajectories extracted within the scope of the study were recorded as two-dimensional images. This image, which consists of combining orbital points, creates distinctive patterns according to different planes.

3. Experimental Results and Discussion

In this study, the dataset used to train the network was created using ultrasound video images. In the training set of the network, three different expert consensus has been labeled so that the organs will be recognized. In the training and test sets of the network, scans were labeled in three different expert consensus for the recognition of anatomical planes. Fetal anatomical structures such as head, arm, heart and trunk obtained from 2D fetal images in the dataset were recognized using the YOLOv5 network. As a result of the experimental studies, the movements in the ultrasound video image were obtained and the movement points were recorded. In addition, in the training of the YOLOv5s network for recognizing the anatomical structures of the fetus and tracking movements, a total of 1300 ultrasound scans belonging to 4 classes were labeled from ultrasound videos, 500 for the head, 400 for the body, 300 for the heart and 100 for the arm. The performance of the proposed YOLOv5 network was validated on both still images and ultrasound videos. A total of 348 US scans for the validation set were used for four classes, 122 US images for the head, 119 US images for the body, 68 US images for the heart, and 39 US images for the arm, for performance measures of the network on still images.

For the experimental studies carried out in this study, the YOLOv5 repository was cloned and the network model was trained on Google Colab. Within the scope of the study, the free Tesla T4 1.5GB GPU provided in the Google Colab system was used. The open original library (Deep-SORT, 2021) was used for the Deep-SORT algorithm used for object tracking. The plots of Box Loss, Class Loss and Object Loss values according to the number of epochs during the training of the YOLOv5 network with 100 epochs is denoted in Figure 2. It is seen that all three loss values are quite low.

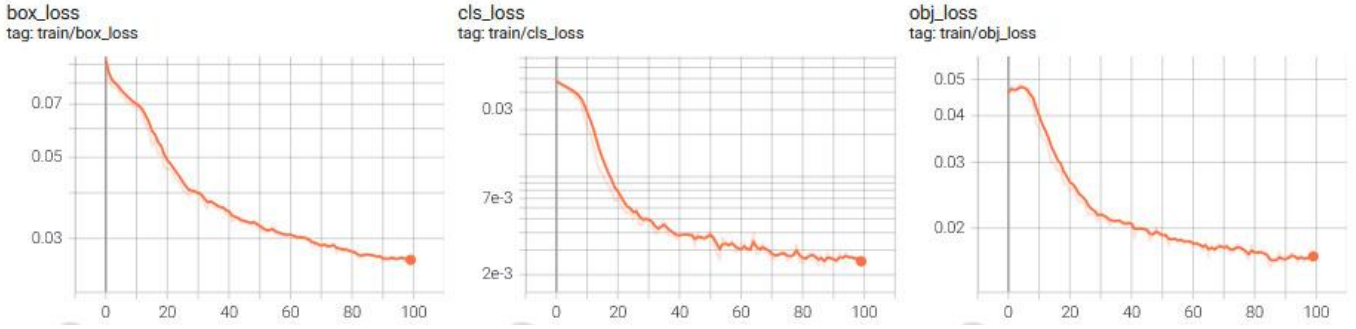


Figure 2. Plots of Box Loss, Class Loss and Object Loss values according to the epoch number during the training of the YOLOv5 network

In this study, the performance of the YOLOv5 network proposed on the validation set for the recognition of anatomical structures in the fetus was measured with the Precision (P), Recall (R) and F1-score (F1) metrics given in Eq. (1), Eq. (2) and Eq. (3). Precision is the rate of how many of the positively predicted objects are actually positive, while the rate of how many of the positively predicted objects are actually positively predicted is the Recall criterion. In the validation set, the training of the YOLOv5 network with 100 epochs, the plots of the mean average precision (mAP), the P and R metrics are indicated in Figure 3. Here, the threshold value of 0.5 for mAP indicates the

overlap of the bounding boxes with the true value. In Figure 3, plots of the box loss, objectness loss and classification loss according to the epoch number in the training are also presented. Box loss represents the loss as the bounding box is plotted as tightly, objectness loss represents the loss calculated for the probability that a region contains an object regardless of the object categories, and classification loss represents the loss related to how accurately the bounding box and its class are determined. As training progresses, the bounding boxes are calculated more tightly and the box loss approaches zero. Also, the classification loss approaches zero as the training progresses.

$$\text{Precision (P)} = \frac{TP}{TP+FP} \tag{1}$$

$$\text{Recall (R)} = \frac{TP}{TP+FN} \tag{2}$$

$$\text{F1 - score (F1)} = 2 \times \frac{P \times R}{P+R} \tag{3}$$

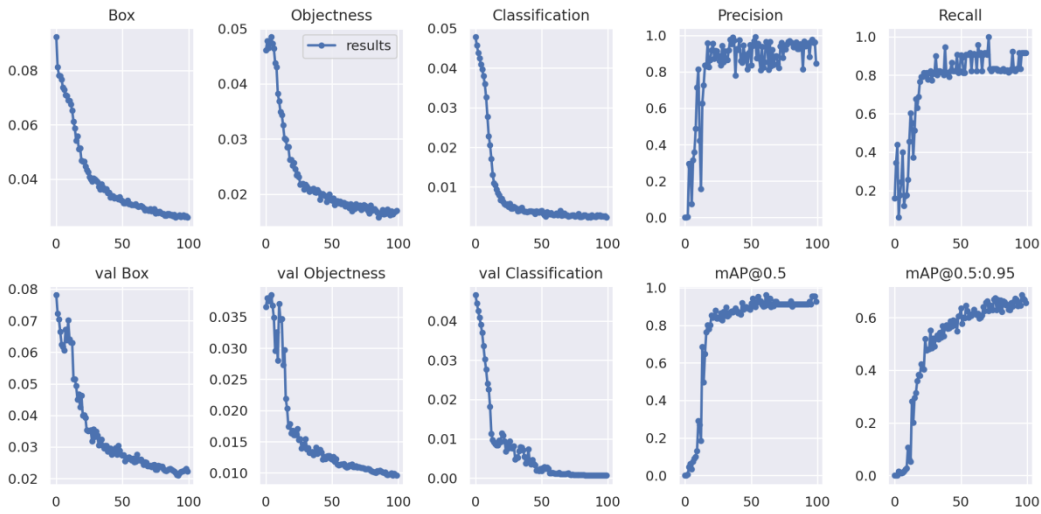


Figure 3. Assessment of box loss, objectness loss, classification loss, P, R and mAP measures during the training epochs in YOLOv5 network for validation set

The results of mAP, P, R and F1 performance metrics obtained in the performance measurements achieved with the validation cluster with 348 scans for the recognition of anatomical structures and organs (such as head, body, heart and arm) in the fetus are presented in Table 1. Also in this table, the confusion matrix values TP, FP and FN values are given for each

plane. In a threshold value of 0.50 for mAP, the mAP scores were obtained as 0.97, 0.91, 0.82 and 0.89 for fetal head, body, heart and arm anatomical structures, respectively. Likewise, F1 scores for fetal head, body, heart and arm anatomical structures were 95.04%, 94.42%, 88.31% and 83.23%, respectively.

Table 1. Comparison of performance metric results such as mAP, P, R and F1 obtained with the YOLOv5 model for validation set

Anatomical plane in fetus	mAP (@0.50)	P (%)	R (%)	F1 (%)	TP	FP	FN
Head	0.97	95.83	94.26	95.04	115	5	7
Body	0.91	96.49	92.44	94.42	110	4	9
Heart	0.82	79.07	100.0	88.31	68	18	0
Arm	0.89	84.21	82.05	83.23	32	6	7

Figure 4 denotes the visual results of plane recognition with the YOLOv5-based model proposed in this study for fetal head, body, heart and arm anatomical structures for some scans in the fetal US dataset. Both still image and US videos show successful

recognition of fetal organs using the proposed YOLOv5-based method. In addition, here, detection scores outside the object recognition bounding boxes in US scans are also specified.

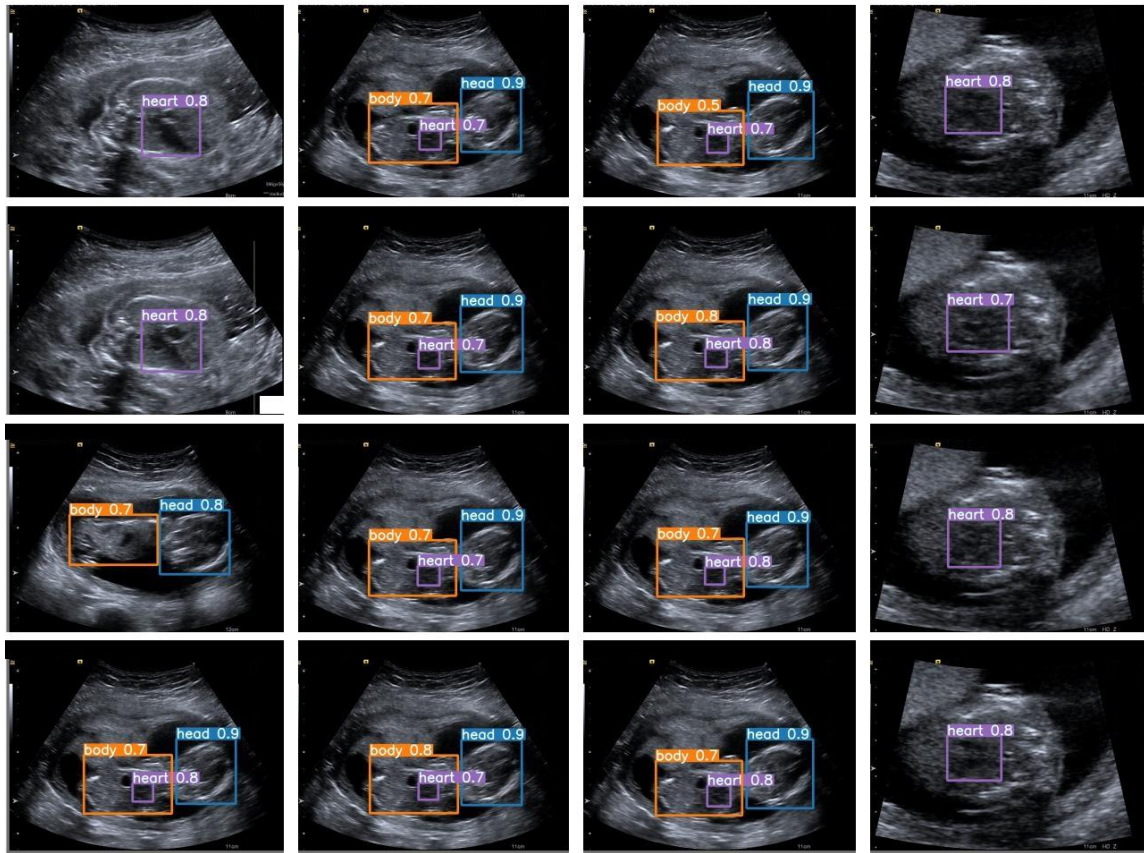
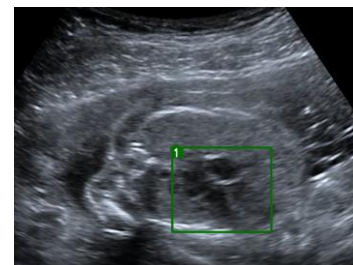
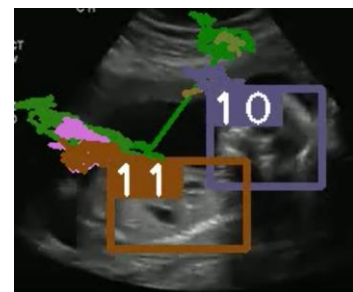


Figure 4. Visual results of the recognition of fetal anatomical planes in US scan using the proposed YOLOv5-based method

In this study, the motions of the planes recognized by the YOLOv5 network in US videos are tracked using the Deep-SORT algorithm and their motion trajectories are extracted. In relation to this situation, in Figure 5(a), the recognition of a plane in US video using the YOLOv5 algorithm and its bounding box are detailed. Moreover, Figure 5(b) shows the tracking of the movements of planes identified with IDs 10 and 11 during the US video using the Deep-SORT algorithm and the extraction of motion trajectories. In the study, instead of the center of the fetus planes, the upper left corner of the bounding box was taken as position information, and the trajectory information of the detected planes was recorded. In our study, trajectory information of different videos obtained from fetal organs was recorded and the characteristics of the movements of the organs were evaluated. In addition, the trajectories of the movements of the head and body anatomical structures of the fetus and the 2D representation of these trajectory patterns in a US video are indicated in Figure 6 and Figure 7, respectively. From this point of view, it is evaluated that fetal movements can be classified with a model that can recognize these trajectories of movements.



(a) Fetal plan in US video



(b) Tracking of the recognized plane on US video

Figure 5. (a) Recognition of the fetal plane on US video using the proposed YOLOv5 algorithm, (b) Tracking motion trajectory of the recognized organ in US video using the Deep-SRT algorithm

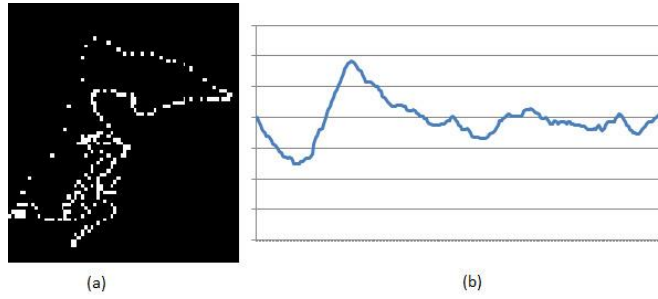


Figure 6. (a) Movement trajectory of the fetal head plane and (b) the 2D plane representation of the trajectory of motion

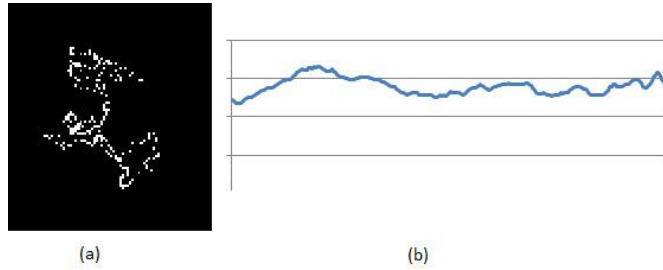


Figure 7. (a) Movement trajectory of the fetal body plane and (b) the 2D plane representation of the trajectory of motion

4. Conclusions

In this study, a deep learning-based approach was proposed to recognize some fetal anatomical structures such as head, body, arm and heart from fetal ultrasound and to detect their movement trajectories. High-performance recognition of fetal planes was achieved by using the YOLOv5 network in the dataset created from videos containing fetal movements. In 0.50 threshold value for mAP, F1 scores for fetal head, body, heart and arm anatomical structures were 95.04%, 94.42%, 88.31% and 83.23%, respectively. In addition, fetal planes were identified using the deep-SORT algorithm, and their motion trajectories were extracted by tracking the movements of the organs on US videos.

In this study, it was observed that the motion trajectory information formed different patterns according to the planes. If enough video images with healthy and unhealthy movements are obtained, it is considered that it may be possible to classify these movements from trajectory images. Also, when each object in an image is named with a different identity, the motion trajectories of each object can be recorded. Obtaining the trajectories of fetal movements may allow the recognition of different movement types.

In the study, it was observed that the raw trajectory data differed for different planes. Therefore, the trajectory data to be used for training the network has to belong to the correct planes. While determining the trajectories, the movements of the ultrasound device can be ignored and real plane movements can be obtained. The movements of the trajectory data to be used for training purposes can be taken with a certain threshold value and position changes caused by the movement of the ultrasound device can not be included in the plane movement. Obtaining the trajectory of motion in ultrasound videos poses a number of challenges. There is no fixed camera angle as in video footage. Due to the movements of the ultrasound device, the position of the image on the screen also changes. This can cause the object on the screen to appear as if it is moving. To overcome this problem, a pixel threshold can be determined for each step of the

motion, which can change. In the study, this situation was not taken into account while extracting the motion trajectories. While evaluating the trajectories, a pixel threshold value for motion can be determined and it can be ensured that plane movements are actually obtained.

While object recognition is performed in ultrasound videos, there may be problems in identifying organs. It is possible that the head image that appears from different angles is marked with a different identity in the deep-SORT algorithm. In order to overcome this situation, while collecting orbital information, determining and addressing individual plane movements can provide meaningful results.

In the following stages of the study, it is aimed to diversify the US dataset. As the number of images in the training set increases, the network can be trained for more planes. Trajectories of detected organ movements can be recorded. Trajectory data can be preprocessed to reduce the motion effects of the ultrasound device. An image can be created from the trajectory information by using the trajectories of the movements of each plane. A network that can learn the differences in plane movements can be designed by training the images created from the trajectory information with a convolutional network. Recognition and identification of movements can also be achieved with a different architecture, which is more suitable for trajectory data instead of a convolutional neural network. In the following stages, a model that can recognize abnormal movements can be created. The movements of the heart plane can be obtained by performing a more detailed labeling. It is possible to work on obtaining the heart graph from the movement pattern.

Acknowledgements

The authors of the study thank Evliya Çelebi Training and Research Hospital of Kütahya Health Sciences University for providing the fetal US dataset. We would also like to express our endless gratitude to Professor Huiyu Zhou from the University

of Leicester, who shared his experiences and contributed to this study.

References

- Ahmed, M., & Noble, J. A. (2016). Fetal ultrasound image classification using a bag-of-words model trained on sonographers' eye movements. *Procedia Computer Science*, 90, 157-162.
- Bai, Y. (2016). *Object tracking & fetal signal monitoring*: Southern Illinois University at Carbondale.
- Baumgartner, C. F., Kamnitsas, K., Matthew, J., Fletcher, T. P., Smith, S., Koch, L. M., Kainz, B., & Rueckert, D. (2017). SonoNet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE transactions on medical imaging*, 36(11), 2204-2215.
- Bewley, A., Ge, Z., Ott, L., Ramos, F., & Upcroft, B. (2016). *Simple online and realtime tracking*. Paper presented at the 2016 IEEE international conference on image processing (ICIP). pp. 3464-3468.
- Carneiro, G., Georgescu, B., Good, S., & Comaniciu, D. (2008). Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree. *IEEE transactions on medical imaging*, 27(9), 1342-1355.
- Deep-SORT. (2021). Deep-SORT Algorithm. Available online: https://github.com/nwojke/deep_sort
- Deepika, P., Suresh, R., & Pabitha, P. (2021). Defending Against Child Death: Deep learning-based diagnosis method for abnormal identification of fetus ultrasound Images. *Computational Intelligence*, 37(1), 128-154.
- Fiorentino, M. C., Moccia, S., Capparuccini, M., Giamberini, S., & Frontoni, E. (2021). A regression framework to head-circumference delineation from US fetal images. *Computer methods and programs in biomedicine*, 198, 105771.
- Gao, Y., Maraci, M. A., & Noble, J. A. (2016). *Describing ultrasound video content using deep convolutional neural networks*. Paper presented at the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 787-790.
- Ishikawa, G., Xu, R., Ohya, J., & Iwata, H. (2019). *Detecting a Fetus in Ultrasound Images using Grad CAM and Locating the Fetus in the Uterus*. Paper presented at the ICPRAM. pp. 181-189.
- Jocher, G. (2020). YOLOv5. Available online: <https://github.com/ultralytics/yolov5>
- Kumar, A., Sridar, P., Quinton, A., Kumar, R. K., Feng, D., Nanan, R., & Kim, J. (2016). *Plane identification in fetal ultrasound images using saliency maps and convolutional neural networks*. Paper presented at the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 791-794.
- Lei, B., Zhuo, L., Chen, S., Li, S., Ni, D., & Wang, T. (2014). *Automatic recognition of fetal standard plane in ultrasound image*. Paper presented at the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI). pp. 85-88.
- Liu, W., Wang, Z., Zhou, B., Yang, S., & Gong, Z. (2021). *Real-time Signal Light Detection based on Yolov5 for Railway*. Paper presented at the IOP Conference Series: Earth and Environmental Science. pp. 042069.
- Looney, P., Stevenson, G. N., Nicolaides, K. H., Plasencia, W., Molloholli, M., Natsis, S., & Collins, S. L. (2017). *Automatic 3D ultrasound segmentation of the first trimester placenta using deep learning*. Paper presented at the 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). pp. 279-282.
- Malathi, G., & Shanthi, V. (2009). *Wavelet based features for ultrasound placenta images classification*. Paper presented at the 2009 Second International Conference on Emerging Trends in Engineering & Technology. pp. 341-345.
- Malta, A., Mendes, M., & Farinha, T. (2021). Augmented Reality Maintenance Assistant Using YOLOv5. *Applied Sciences*, 11(11), 4758.
- Ravishankar, H., Prabhu, S. M., Vaidya, V., & Singhal, N. (2016). *Hybrid approach for automatic segmentation of fetal abdomen from ultrasound images using deep learning*. Paper presented at the 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 779-782.
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). *You only look once: Unified, real-time object detection*. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 779-788.
- Salomon, L. J., Alfirevic, Z., Berghella, V., Bilardo, C., Hernandez-Andrade, E., Johnsen, S., Kalache, K., Leung, K. Y., Malinger, G., & Munoz, H. (2011). Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound in Obstetrics & Gynecology*, 37(1), 116-126.
- Sinclair, M., Baumgartner, C. F., Matthew, J., Bai, W., Martinez, J. C., Li, Y., Smith, S., Knight, C. L., Kainz, B., & Hajnal, J. (2018). *Human-level performance on automatic head biometrics in fetal ultrasound using fully convolutional neural networks*. Paper presented at the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 714-717.
- Sobhaninia, Z., Rafiei, S., Emami, A., Karimi, N., Najarian, K., Samavi, S., & Soroushmehr, S. R. (2019). *Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning*. Paper presented at the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 6545-6548.
- Soviany, P., & Ionescu, R. T. (2018). *Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction*. Paper presented at the 2018 20th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC). pp. 209-214.
- van den Heuvel, T. L., Petros, H., Santini, S., de Korte, C. L., & van Ginneken, B. (2019). Automated fetal head detection and circumference estimation from free-hand ultrasound sweeps using deep learning in resource-limited countries. *Ultrasound in medicine & biology*, 45(3), 773-785.
- Velazquez, M. D., & Rayburn, W. F. (2002). Antenatal evaluation of the fetus using fetal movement monitoring. *Clinical obstetrics and gynecology*, 45(4), 993-1004.
- Wang, C.-Y., Liao, H.-Y. M., Wu, Y.-H., Chen, P.-Y., Hsieh, J.-W., & Yeh, I.-H. (2020). *CSPNet: A new backbone that can enhance learning capability of CNN*. Paper presented at the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. pp. 390-391.
- Whitehead, C. L., Cohen, N., Visser, G. H., & Farine, D. (2020). Are increased fetal movements always reassuring? *The*

- Journal of Maternal-Fetal & Neonatal Medicine*, 33(21), 3713-3718.
- Wojke, N., Bewley, A., & Paulus, D. (2017). *Simple online and realtime tracking with a deep association metric*. Paper presented at the 2017 IEEE international conference on image processing (ICIP). pp. 3645-3649.
- Wróbel, J., Kupka, T., Horoba, K., Matonia, A., Roj, D., & Jeżewski, J. (2014). Automated detection of fetal movements in Doppler ultrasound signals versus maternal perception. *Journal of Medical Informatics & Technologies*, 23.
- Yan, B., Fan, P., Lei, X., Liu, Z., & Yang, F. (2021). A Real-Time Apple Targets Detection Method for Picking Robot Based on Improved YOLOv5. *Remote Sensing*, 13(9), 1619.
- Yaqub, M., Napolitano, R., Ioannou, C., Papageorghiou, A., & Noble, J. A. (2012). *Automatic detection of local fetal brain structures in ultrasound images*. Paper presented at the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI). pp. 1555-1558.
- You, J., Li, Q., Guo, Z., & Zhao, R. (2017). *Smart fetal monitoring*. Paper presented at the International Conference on Information Science and Applications. pp. 494-503.
- Yu, Z., Ni, D., Chen, S., Li, S., Wang, T., & Lei, B. (2016). *Fetal facial standard plane recognition via very deep convolutional networks*. Paper presented at the 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). pp. 627-630.