Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 221 (2023) 1–8

Procedia Computer Science

www.elsevier.com/locate/procedia

10th International Conference on Information Technology and Quantitative Management

Does Peer Pressure Have The Same Effect On Intelligent Decision Support System As It Does On Human Decision Systems? Case Study: Fetal Ultrasound Movies

Smaranda Belciug[a]*

[a]*University of Craiova, Craiova, 200585, Romania*

Abstract

The aim of this paper is to explore whether the peer pressure environment makes Artificial Intelligence machines behave just like humans do. We propose a two-step multi-participant intelligent decision support system that uses a committee of deep learning algorithms and soft voting classifier to classify fetal ultrasound morphology images. The formed intelligent decision support system is subjected to peer pressure when used to classify fetal ultrasound movies. The results show that in 82% cases the peer pressure environment makes the intelligent decision system take the correct decision.

Keywords: deep learning; multi-participant intelligent decision support system; soft voting classifier; fetal morphology; ultrasound movie.

1. Introduction

We live in a world where peer pressure makes us do impossible things, things that we might never considering doing. Peer pressure is part of our daily routine, even if we like it or not. We take into consideration the opinion of people that surround us, people that pass by us on the street, we make decisions trying to predict their reaction, and our response to their reaction. We are social beings, and it is natural to behave this way. But can we consider peer pressure some sort of multi-participant collaborative-decision making process? Are humans prone to always make individual decisions based on the *a priori* decisions of their fellows? What about Intelligence Decision Support

10.1016/j.procs.2023.07.001

Systems (IDSS)? Are IDSS affected by the peer pressure? And, ultimately, is peer pressure a good or a bad thing when it comes to IDSS?

As all things in life, it all depends on the perspective of the problem. Imagine you are in a room, and you take a look around you. Inevitably your next move will depend on the way the objects, people, furniture, etc., are represented in your mind. But what if you were to climb onto a chair or a table? Will the perspective be the same, will the room and its components look identical? No. Hence you will adapt your moves according to the new representation. Since, not all people can stay in the same place, at the same time, it is natural that they have a different perspective of the same issue, and the same issue might affect differently the participants. Having this in mind, is it correct to make a decision based only on one perspective? Isn't it better to relate on peer pressure? Should we let our future plan be created by a multi-participant decision system? If the problem on site is not a crucial one, maybe relying on just our opinion is not such of a big a deal, but if we were to consider a problem that is bigger than us, should we still ponder whether to take into account other people's ideas?

When it comes to medical decision making, things are never easy. There are discrepancies between multiple doctors, and now with the arise of Artificial Intelligence in this area, there are discrepancies between doctors, data scientists, and IDSS. Should a medical decision be based only one doctor, only on an IDSS, or should we deepen the problem and merge several doctors with several IDSS, and base our medical diagnosis or treatment on peer pressure?

In this paper, we are going to discuss an important area in the Obstetrics medical field, the fetal morphology scan. The International Society of Ultrasound in Obstetrics and Gynecology, The Fetal Medicine Foundation, and other national and international societies recommend women to attend their second trimester ultrasound to determine whether their fetus suffers from congenital anomalies or not. If the ultrasound is read accurately, a detailed discussion with the future-to-be parents takes place. This discussion is crucial, and it hits the following hot topics: long term morbidity, mortality, quality of life, and procedural risks. The correct interpretation of a fetal morphology scan is observer dependent, therefore, the less experienced a sonographer is, the more anomalies could be missed. The discrepancies between pre- and postnatal diagnosis range between 27.5% and 96% in terms of sensitivity, [1]. In the first 28 days after birth, approximately 300 000 newborns die because of a missed or incorrect diagnoses of congenital anomalies. Romania is placed on top in this statistic, [2]. Some of these anomalies can be treated or at least controlled if spotted at the right time during an ultrasound. However, some cannot, and these anomalies leave 3.2 million children disabled for life, [3, 4]. Congenital anomalies can be detected during the second trimester fetal morphology scan. This is not a routine procedure since it needs to be performed by a skilled sonographer. The detection rate for an experience sonographer is of only 52% accuracy, while for an unexperienced one is of 32.5% accuracy.

In this paper, we propose a novel theory of how to improve the diagnosis using a multi-participant decision system that is based on different artificial intelligence methods, doctors, and data scientists that form the peer pressure that makes the final IDSS take a correct decision. This advanced IDSS has the capacity to gather and analyze data, communicate with other systems, take into consideration their opinion on a certain matter, learn from experience, and adapt according to the newer cases. Technically, our aim is to foster a cross-fertilization of fetal morphology, image processing, and Artificial Intelligent methods that rise above the boundaries of the disciplines involved, and lead to new impactful methods that can and will assist medical practice and discovery.

There are several studies that regard the use of Artificial Intelligence on fetal morphology scans. Different deep learning algorithms have been applied to differentiate between fetal ultrasound images, [5-10] or segment the fetal brain and lungs, [11, 12]. None of these studies, do not cover a multi-participant IDSS, to make the final decision. The aim of this study is to present how a multi-participant IDSS can create the peer pressure necessary for a IDSS to classify correctly fetal morphology movies, being previously trained on ultrasound images. The paper is organized as follows: section 2 presents the concept of a multi-participant decision system, together with the Artificial Intelligence algorithms that are used to form it. Section 3 presents the dataset that has been used for training purposes, and also depicts the design and application of the multi-participant IDSS for video classification. Section 4 covers the obtained results, and the paper ends with the conclusions in section 5.

2. Multi-participant IDSS or how does peer pressure influence the response of an IDSS

Peer pressure is defined by Cambridge Dictionary as "the strong influence of a group on member of that groups to behave as everyone else does". Is peer pressure a form of collaboration or cooperation? Cooperation means "the action or process of working together to the same end", while collaboration means "the action of working with someone to produce something". Both cooperation and collaboration in habitats where individuals are self-interested is difficult to achieve, [13]. When each individual works rationally in her/his/their self-interest, will most likely cause a drain of a common resource to the detriment of other individuals. But if we place an individual into a group that already collaborates, and acts accordingly in a certain manner, then that individual will not resist the peer pressure and will start acting in the same manner as the rest.

The multi-participant decision system contains the subsequent phases, [14, 15]:

- In the *preparation* phase the problem is defined together with its domain, state-of-the-art, constraints, criteria, and decision unit.
- The *collective understanding* phase comprises finding the common route to understand the issue that needs to be resolved, and how to implement the decision process.
- The *solution* phase consists of designing the decision-making process and covers its alternatives.
- The *negotiation and confrontation* phase consists of the participants discussions. Everyone presents his/her/their viewpoint and tries to convince the others to accept and support their decision.
- The *decision* phase establishes the final decision.
- The *monitoring* phase consists of monitoring the decision process to ensure that it goes smoothly, and no problems appear. In the case, when problems still appear, they are caught and fixed at an early stage.

Our multi-participant decision system has two phases which can be viewed as specific contribution to the Healthcare 4.0 movement inspired from the Industry 4.0 movement, [16, 17]. In the first phase the participants are 3 different deep learning (DL) algorithms (ResNet50, EfficientNetV2S, and MobileNet3Larger, [19-20]) that are trained on the ultrasound images dataset. The same classifier was used for all 3 DL models, having the following architecture: GlobalAveragePooling2D, Dense (1024, activation = ReLU), Dense (softmax). As for the optimizer we have used adam, and as the loss function we have used the categorical crossentropy. The training process was done after performing power analysis. Therefore, we have estimated the appropriate sample size for obtaining a default statistical power goal $P \geq 95\%$ and type I error $\alpha = 0.05.$ to be 50 independent runs in a complete 10-fold cross-validation cycle. After the training process is done, we proceed with the testing phase, and we record the performances of each DL in terms of average accuracy and standard deviation. Having the three independent samples, we have built a multi-participant voting system using the soft voting classifier.

Soft voting classifier takes the probabilities of each DLs' prediction, merges them, and chooses as the final decision of the system the prediction that gives the highest total probability. In this way, the multi-participant mechanism will balance the potential weakness of each algorithm:

$$y = \underset{i}{\operatorname{argmax}} \frac{1}{N_{classifiers}} \sum_{i=1}^{N_{classifiers}} (p_1, p_2, \dots, p_i).$$

where p_i is the probability of DL i.

The summarized decision system is:

1. Choose different deep learning models.
2. Train the above models on a specific dataset that contains images.
3. Record the performances obtained by the models on the testing set.
4. Apply the soft voting formula using each models' prediction.
5. Label a movie using the labels of the preceding 30 frames.

After this stage is over, we have an IDSS that can classify ultrasound images of an abdomen fetal morphology scan. In order to make the IDSS classify ultrasound movies, we shall appeal to the concept of "peer pressure". Technically, we split the ultrasound movie into consecutive frames. The IDSS created at step 1 is applied on 30 consecutive frames. We have chosen the number 30 in a meta-heuristic manner, trying several runs with different values that ranged from 15 to 50. The best result was achieved by the value 30. Each one of the 30 frames has assigned a label by the IDSS. The whole ultrasound movie will be labeled considering the labels of the 30 frames. Concretely, each frame will give a vote, its label, and the label that has the majority of votes, will be the peer pressure that will give the following frames the same label, and thus will label the entire ultrasound movie. In real time, the ultrasound waves are sent back to the transducer and thus a moving picture of different anatomical parts of the fetus is created. As the mother breathes, and the fetus moves, in a movie that regards a view-plane it is possible that multiple images from other different view planes to appear. So, for us to diagnose congenital anomalies, we must establish which is the predominant view plane we are in. Hence, the peer pressure appears in the picture: no matter what view-plane the system detects (may that view plane be correct), it must take into account the majority of votes of the preceding frames.

3. Fetal morphology dataset

The dataset used in this study comes from a prospective cohort study that takes place in the maternity ward from the University Emergency County Hospital of Craiova. The eligible patients admitted themselves at the hospital's Prenatal Unit for their second trimester morphology scan. They have been informed and invited to take part of the research. The scans were acquired using Voluson 730 Pro (GE Medical Systems, Zipf, Austria), and Logic e (GE Healthcare, China US machines) with 2-5-MHz, 4-8-Mhz, and 5-9-MHz curvilinear transducers. We have collected the ultrasound movies from 100 patients. The dataset has 1388 images divided in 9 decision classes that cover the fetal abdomen view planes and has the following distribution: 3 vessels plus bladder (308 images), abdominal circumference (172 images), anteroposterior kidney (174 images), bladder (55 images), echogenic (162 images), gallbladder (115 images), kidney longitudinal renal (199 images), cord insertion sagittal (81 images), and transverse cord insertion (122 images). Figure 1 presents a sample image from each decision class.

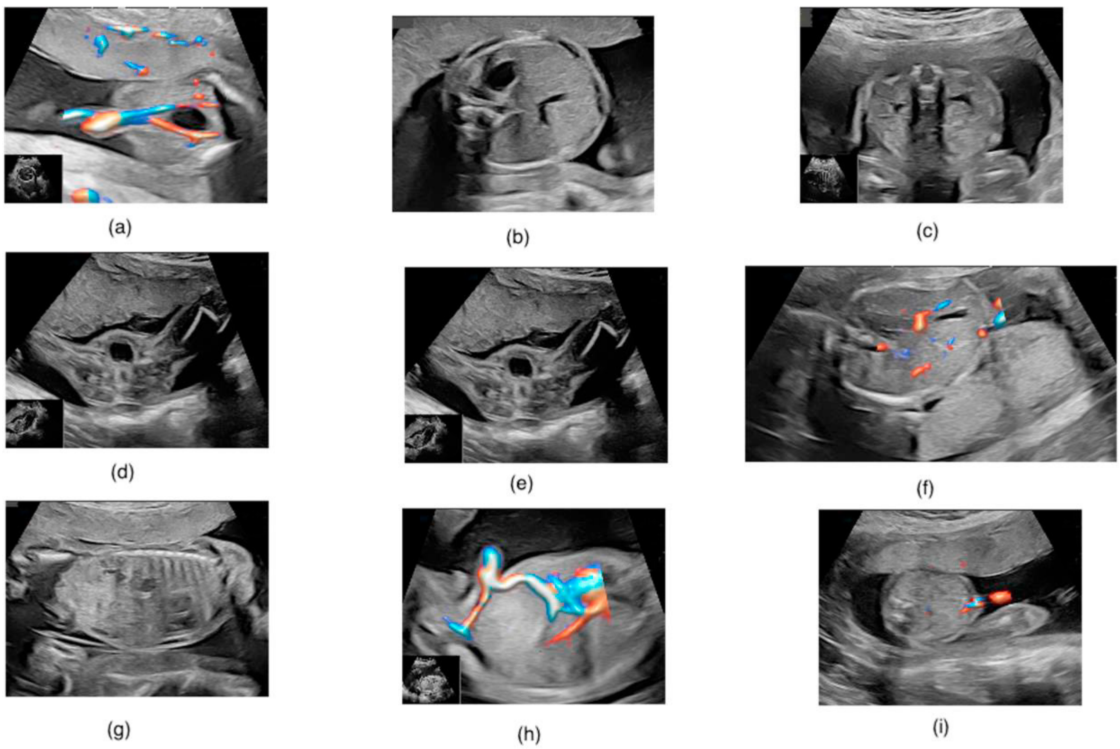

Fig. 1. (a) 3 vessels + bladder; (b) abdominal circumference; (c) anteroposterior kidney; (d) bladder; (e) echogenic; (f) gallbladder; (h) kidney longitudinal renal; (i) cord insertion sagittal; (j) transverse cord insertion

We have used a data generator to increase the sample data. The different transformations were applied to each image using the subsequent values: shear rage 0.2, rotation range 20, zoom range 0.2, brightness range $\in [0.7, 1.4]$, width shift range = 0.1, height shift range = 0.1. The images were resized to be 224×224 px.

We have preprocessed all the images so that the text and other artefact to be removed. CV2 and Keras-OCR have been deployed to detect, remove and inpaint these areas using the surrounding pixels. Figure 2 presents a sample image before and after preprocessing.

Fig. 2. (a) unprocessed picture; (b) preprocessed image

4. Results

To obtain reliable results and avoid any potential information leakage, we have split the dataset into training and testing, and held out one patient's ultrasound movie for testing purposes. As stated in the previous section, we have run the algorithms independently for 50 runs each in a complete cross-validation cycle. The results in terms of average accuracy (ACA) and standard deviation are presented in table 1.

Table 1. Performances in terms of ACA and SD

Model	ACA	SD
ResNet50	79.23	2.99
EfficientNetV2S	81.34	1.60
MobileNetV3Large	80.12	1.87

From table 1, we can see the that the performances of the three models are good and have small standard deviation, which shows that they are robust and stable.

Having obtained the above results, we have proceeded in building the multi-participant IDSS based on the soft voting classifier. The performance of the newly formed IDSS reached 82.19% ACA. The next step in our experiment was to use the peer pressure of the previously labeled frames to classify the whole ultrasound movie. In figure 3 we see how the multi-participant two-step IDSS sets the label for the entire movie that regards the view plane abdominal circumference

From figure 3, we can see that the IDSS identified 27 frames as being abdominal circumference, 2 as being gallbladder and 1 as being bladder. After confronting with the medical professional, it was clear that indeed the algorithm did not make a mistake when classifying 3 frames differently, because indeed in those frames, due to the fetus' movement and mother's breathing the view-plane switched from abdominal circumference to gallbladder and bladder, respectively. Nevertheless, our proposed "peer-pressure" approach was able to correct the IDSS. This is the first step in detecting congenital anomalies, the following step representing finding the appropriate anatomical structures in the respective frame.

Fig. 3. IDSS classifies the abdominal circumference view plane

Fig. 4. IDSS classifies the sagittal cord insertion view plane

In figure 4 we have a more complex example. Here we are testing the IDSS to see whether it can correctly classify the sagittal cord insertion view plane. From this example we can see that the IDSS classified 20 images as being sagittal cord insertion, 9 as being echogenic, and just 1 as being abdominal circumference. This is an especially hard example since the only differences between the view planes sagittal cord insertion and echogenic is the presence of the umbilical cord. Indeed, the movement of the fetus makes the cord disappear in some frames, hence the echogenic class is identified. Nevertheless, the multi-participant two-step IDSS is able to correct this honest mistake.

Fig. 5. IDSS tries to classify longitudinal kidney

In figure 5, we have an example of peer pressure gone wrong. The IDSS tries to classify the longitudinal kidney plane but fails because of peer pressure. The IDSS classified the ultrasound movie as being anteroposterior kidney view plane. Taking into account that the overall testing performance is around 82%, these situations are bound to happen. One way to resolve this issue is to increase the number of images in the training dataset.

An issue that remains to be resolved is how peer pressure will resolve differentiating correctly the view planes, when the sonographer moves the transducer. The system might be confused when computing the voting due to having multiple frames from the precedent class. Further research in this direction must be done. One potential alternative might be the use of 3D neural networks.

5. Conclusions

In this study, we proposed a new approach of the multi-participant IDSS, which is based on the peer pressure of the surrounding individuals. We have wanted to see if peer pressure works the same for machines, as it does for human beings. So far, the results seem promising. However, the drawbacks of peer pressure on human beings, can translate into drawbacks in peer pressure for machines. If us, people, surround ourselves with negative examples, we will soon find ourselves making wrong decisions. Group intelligence can lead to improved decision making, problem solving, innovation, efficiency, creativity, etc. Nevertheless, in certain situations the "wisdom of crowds" can lead to potential limitations such as the lack of diversity and variation, and even the appearance of a compromise, rather than the optimal solution.

The same principals apply to machines also. To correct these situations, we must train our DLs on larger images datasets. Future work will include deepening the experiments by increasing the size of the dataset, using other voting classifier methods, and other DLs.

Acknowledgements

This work was supported by a grant of the Ministry of Research Innovation and Digitization, CNCS-UEFISCDI, project number PN-III-PCE-2021-0057, within PNCDI III.

References

[1] Salomon, L. et al (2008) "A score-based method for quality control of fetal images at routing second trimester ultrasound examination" *Prenat Diag*, **28** (**9**), 822-827.

[2] https://data.unicef.org/country/rou

[3] Kinser-Ovaskainen, A. et al. (2020) "European Monitoring of Congenital Anomalies: JRC EUROCAT Report on Statistical Monitoring of Congenital Anomalies (2008-2017), EUR 30158 EN, Publication Office of the European Union, Luxembourg, doi:10.2760/65886.

[4] Lobo, I., Zhaurova, K. (2008) "Birth defects: causes and statistics. *Nature Education*, **1** (**1**), 18.

[5] Phillip, M. et al. (2019) "Convolutional Neural Networks for automated fetal cardiac assessment using 4D B-Mode ultrasound, *IEEE 16th Int Symp Biomed Imag*.

[6] Matsuoka, R, Komatsu, M, et l. (2019) "A novel deep learning based system for fetal cardiac screening" *Ultras Obstet Gynecol*, https://doi.org/10.1002.uog.20945.

[7] Komatsu, R, Matsuoka, R., et al. (2019) "Novel AI-guided ultrasound screening system for fetal heart can demonstrate finding in timeline diagram" *Ultras Obstet Gynecol*, https://doi.org/10.1002/uog.20796.

[8] Burgos-Artizzu, X.P., et a. (2020) "FETAL_PLANES_DB: common maternal-fetal ultrasound images" *Nature Scientific Reports*, **19**, 10200.

[9] Belciug, S. (2022) "Learning deep neural networks' architectures using differential evolution. Case study: medical imaging processing". *Computers in Biology and Medicine*, **146**, 105623.

[10] Ivanescu, R. et al, (2022) "Evolutionary computation paradigm to determine deep neural networks architectures. *Int J Comp Comm Control*, **17** (**5**), 4886, https://doi.org/10.15837/ijccc.2022.5.4886.

[11] Namburete, A., et al. (2018) "Fully automated alignment of 3D fetal brain ultrasound to a canonical reference space using multi-task learning. *Med Image Anal*, **46**, 1-14.

[12] Torrents-Barrena, J. et al. (2019) "Assessment of radiomics and deep learning for the segmentation of fetal and maternal anatomy in magnetic resonance imaging and ultrasound, *Acad Radiol*, **19**, 30575-30576.

[13] Filip, F.G., Zamfirescu, C.-B., Ciurea, C. (2018) "Computer-Supported Collaborative Decision-Making", https://doi.org/10.1007/978-3-319-47221-8, Springer.

[14] Filip, F.G. (2022) "Collaborative Decision-making: concepts and supporting information and communication technology tools and systems" *Int J Comp Comm Control*, **17** (**2**), 4732.

[15] Konate, J., Gueye, A., Zarate, P., Camillieri, G. (2023) "Collaborative decision-making: a proposal of a semi-automatic facilitation based on an ontology", *International Journakl of Information Technology & Decision Making,* **22** (**1**), 447-470.

[16] Wehde, M. (2019) "Healthcare 4.0", *IEEE Engineering Management Review*, **44**, (**3**), 24081, 10.1109/EMR.2019.2930702.

[17] Tortella, G.L., et al. (2020) "Healthcare 4.0: trends, challenges and research directions, *Production Planning & Control*, **31**, (**15**), 1245-1260, 10.1080/09537287.2019.1702226.

[18] He, K., Zhang, X., Ren, S., Sun, J. (2015) "Deep Residual Learning for Image Recognition", arxiv.org/abs/1512.03385

[19] Tan, M., Le, Q. (2021) "EfficientNetV2: smaller models and faster training", arxiv.org/abs/2104.00298

[20] Howard, A., Sandler, M., et al. (2019) "Searching for MobileNetV3", arxiv.org/abs/1905.02244.