

Yasunari Miyagi\*, Toshiyuki Hata, Saori Bouno, Aya Koyanagi and Takahito Miyake

# Recognition of facial expression of fetuses by artificial intelligence (AI)

<https://doi.org/10.1515/jpm-2020-0537>

Received November 17, 2020; accepted December 27, 2020;

published online February 8, 2021

## Abstract

**Objectives:** The development of the artificial intelligence (AI) classifier to recognize fetal facial expressions that are considered as being related to the brain development of fetuses as a retrospective, non-interventional pilot study.

**Methods:** Images of fetal faces with sonography obtained from outpatient pregnant women with a singleton fetus were enrolled in routine conventional practice from 19 to 38 weeks of gestation from January 1, 2020, to September 30, 2020, with completely de-identified data. The images were classified into seven categories, such as eye blinking, mouthing, face without any expression, scowling, smiling, tongue expulsion, and yawning. The category in which the number of fetuses was less than 10 was eliminated before preparation. Next, we created a deep learning AI classifier with the data. Statistical values such as accuracy for the test dataset and the AI confidence score profiles for each category per image for all data were obtained.

**Results:** The number of fetuses/images in the rated categories were 14/147, 23/302, 33/320, 8/55, and 10/72 for eye blinking, mouthing, face without any expression, scowling, and yawning, respectively. The accuracy of the AI fetal facial expression for the entire test data set was 0.985. The accuracy/sensitivity/specificity values were 0.996/0.993/

1.000, 0.992/0.986/1.000, 0.985/1.000/0.979, 0.996/0.888/1.000, and 1.000/1.000/1.000 for the eye blinking, mouthing, face without any expression, scowling categories, and yawning, respectively.

**Conclusions:** The AI classifier has the potential to objectively classify fetal facial expressions. AI can advance fetal brain development research using ultrasound.

**Keywords:** artificial intelligence; deep learning; facial recognition; fetus; machine learning; ultrasonography.

## Introduction

Fetal behaviors such as fetal facial expressions and fetal movements that can be observed by three- or four-dimensional ultrasound have been considered as being related to the development of the fetal brain and central nervous system [1–10]. Kurjak reported a scoring system [11], which was later modified by Stanojevic [12], to assess fetal neurobehavioral development by assessing fetal facial expressions and movements. Fetal facial expressions such as eye blinking, yawning, sucking, mouthing, tongue expulsion, scowling, and smiling can be assessed by four-dimensional ultrasound from the beginning of the second trimester of pregnancy [2, 13]. Emotional behaviors such as smiling at comfortable times, and responses to different stimuli in the womb, are reflected in fetal faces [14] and the frequencies of each facial expression are related at week 13 of gestation. Reissland reported that fetuses of smoking mothers showed higher mouthing rates compared to non-smokers [15]. Therefore, it is important to investigate fetal facial expressions that are considered to be related to fetus brain development. However, there are no standard methods to objectively evaluate fetal facial expressions.

Artificial intelligence (AI) has advanced and entered medicine recently. In various fields of obstetrics and gynecology, studies related to AI have been published such as the estimation of fetal weight [16], the diagnosis of colposcopy [17, 18], the prediction of live births [19–22], the prediction of results of clinical trials [23, 24], the threshold criteria for fibrinogen and fibrin/fibrinogen degradation products in cases of massive hemorrhage during delivery [25], and so on. The well-trained AI classifier, which is a

---

\*Corresponding author: Yasunari Miyagi, MD, PhD, Director, Department of Gynecology, Miyake Ofuku Clinic, 393-1 Ofuku, Minami Ward, Okayama, Okayama Prefecture 701-0204, Japan; Representative, Medical Data Labo, 289-48 Yamasaki, Naka Ward, Okayama, Okayama Prefecture 703-8267, Japan; and Visiting Professor, Department of Gynecologic Oncology, Saitama Medical University International Medical Center, Hidaka, Saitama Prefecture 350-1298, Japan, Phone: +81 86 281 2020, Fax: +81 86 281 7575, address: E-mail: ymiyagi@mac.com

Toshiyuki Hata, Department of Obstetrics and Gynecology, Miyake Clinic, Okayama, Japan; Department of Perinatology and Gynecology, Kagawa University Graduate School of Medicine, Kagawa, Japan  
Saori Bouno and Aya Koyanagi, Department of Obstetrics and Gynecology, Miyake Clinic, Okayama, Japan

Takahito Miyake, Department of Gynecology, Miyake Ofuku Clinic, Okayama, Japan; Department of Obstetrics and Gynecology, Miyake Clinic, Okayama, Japan

computer program that can objectively assess and classify fetal facial expressions, would help to investigate the development of the fetal brain and central nervous system. In this article, we build the original neural network architecture of the AI classifier and demonstrate the feasibility of AI to classify fetal facial expressions as a pilot study.

## Materials and methods

### Patients and data preparation

The study collected images of fetal faces with a four-dimensional ultrasound technique obtained from consecutive singleton pregnancies of outpatients in routine conventional practice from 19 to 38 weeks of gestation with informed consent from all individuals at the Miyake Clinic from January 1, 2020, to September 30, 2020, with completely de-identified data enrolled.

Several minutes after the patients were comfortably lying in the supine position at least one fetal facial still image per fetus was captured and stored using GE Voluson E10 or Voluson E10 BT20 (GE Healthcare Japan). In discussions of a doctor and two sonographers, all images were subsequently classified into seven facial categories; eye blinking, mouthing, face without any expression, scowling, smiling, tongue expulsion, and yawning. The images classified as the supervised dataset were transferred to our AI system at Medical Data Labo offline. This retrospective, non-interventional study was approved by the Institutional Review Board (IRB) of Miyake Clinic (IRB No. 2019-10).

### Fetal face images

Each image of the fetal region of interest is cropped into a square and then saved in size of 100×100 pixels to provide the best accuracy. The image has been de-identified so as not to identify the person.

### AI classifier

The AI classifier that comprised both the convolutional neural network (CNN) [25–31] with L2 regularization [32, 33] to obtain the probabilities of predicting each category of the fetal face was developed. We introduced deep learning for images with an original CNN architecture as shown in Table 1. The CNN in the image comprised 13 layers with a combination of convolutional layers, pooling layers [34–36], flattened layers [37], linear layers [38, 39], and rectified linear unit layers [40, 41], batch normalization layer [42], and softmax layer [43, 44] that presented the probabilities of each category as confidence scores. The category with the highest probability was defined as the AI classification category of each image.

The category, in which the number of fetuses was less than 10, was removed before preparation because it was too short to create AI. Then, all fetus-rated data were divided into test data sets and training data sets at random in a ratio of one to four. Eighty percent of the training data set was used as the AI training data set. The remaining 20% of the data set was defined as the validation data set. In this way, AI training data sets, validation data sets, and non-overlapping test

**Table 1:** The neural network architecture of the classifier for recognizing fetal facial expression.

The sequence number for processing	Layers
1	Convolution layer
2	ReLU layer
3	Pooling layer
4	Convolution layer
5	ReLU layer
6	Pooling layer
7	Flatten layer
8	Linear layer
9	ReLU layer
10	Linear layer
11	Batch normalization layer
12	Linear layer
13	Softmax layer

ReLU, rectified linear unit.

data sets were created. The AI classifier was trained by an AI training data set with concurrent validation by the validation data set, and then the AI classifier was evaluated with the test data set. The training data set is augmented by rotating the images by 0, 90, 180, and 270°, as is often done in the AI classifier process known as data augmentation, because image processing with any degree of rotation can produce images, resulting in different vector data of the same category [21]. The feasibility of the AI classifier was evaluated for the test data set. The statistical values of the test data set are obtained, such as accuracy, sensitivity, specificity, etc. Confidence score profiles were obtained for each category per image for all data.

### Development environment

The development tools and conditions used are: Intel Core i5 with Windows 10 (Redmond, WA, United States), 32 GB (Santa Clara, CA, United States), NVIDIA GeForce GTX 1080 Ti (Santa Clara, CA, United States), and Wolfram Language 12.0 (Wolfram Research, Champaign, IL, United States).

### Statistical analysis

Wolfram Language 12.0 is used for all statistical analyzes. A one-way analysis of variance (ANOVA) test with Sheffe's post hoc test was used. The p-value <0.05 was considered statistically significant.

## Results

The number of fetuses/images was 14/147, 23/302, 33/320, 8/55, 3/16, 2/10, 10/72, and 93/922 for eye blinking, mouthing, face without any expression, scowling, smiling, tongue expulsion, yawning, and all categories, respectively. Then the categories scored in this study were eye blinking, mouthing, face without any expression,

scowling, and yawning. The total number of fetuses/images in the rated categories was 88/896. Maternal age with mean±standard deviation (SD) (range) was 31.5±4.32 (24–39), 33.29±4.50 (26–41), 32.16±5.07 (20–42), 34.25±4.89 (27–42), 35.4±4.27 (31–42), and 32.81±4.75 (20–42) years for eye blinking, mouthing, face without any expression, scowling, yawning, and all categories, respectively.

There were no significant differences for maternal age between the five categories (p=0.36 by two-sided one-way ANOVA test). The mean gestational week±SD (range) was 34.29±2.77 (28.00–38.4), 29.69±3.81 (19.29–36.423), 31.07±2.92 (27.4–36.86), 32.92±3.55 (27.43–36.86), 29.89±4.89 (19.29–38.43), and 31.37±3, 76 (19.29–38.43) weeks for eye blinking, mouthing, face without any expression, scowling, yawning, and all categories, respectively. There were significant differences for weeks of gestation between the five categories (p<0.01 by the two-sided one-way ANOVA test). The weeks of gestation when they blinked and scowled were more advanced than the other categories (p<0.05 according to Sheffe’s post hoc test).

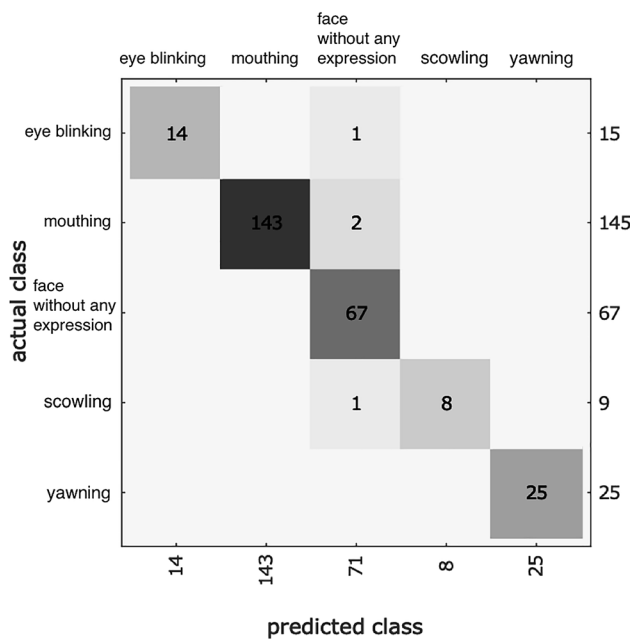
The confusion matrix plot for the test dataset was shown in Figure 1. The correct predictions were located in a diagonal matrix position.

The accuracy of fetal facial expression for the entire test dataset was 0.985. The accuracy values were 0.996, 0.992, 0.985, 0.996, and 1.000 for the eye blinking, mouthing, face without any expression, scowling, and

yawning categories, respectively, as shown in Table 2. The sensitivity values were over 0.933 except for the scowling category. All specificity values were greater than 0.979. All positive and negative predictive values were over 0.943. All F1 scores that are the harmonic mean of the sensitivity value and the positive predictive value were greater than 0.941. All information values [45], which is a multiclass generalization of Youden’s J Statistic [46], were greater than 0.889. The cross-entropy was 12.59. The [mean±SD of confidence scores for eye blinking, mouthing, face without any expression, scowling, and yawning category] was [0.51±0.35, 0.10±0.14, 0.34±0.35, 0.004±0.004 and 0.04±0.08], [0.0003±0.0005, 0.98±0.07, 0.01±0.06, 0.0006±0.0015 and 0.006±0.016], [0.00011±0.00033, 0.001±0.010, 0.997±0.016, 0.00006±0.00017, 0.002±0.008 and 0.0003±0.0004, 0.017±0.032, 0.12±0.22, 0.85±0.23 and 0.011±0.018] and [0.000025±0.000032, 0.0004±0.0016, 0.01±0.07, 0.00010±0.00017 and 0.99±0.07] in the eye blinking, mouthing, face without any expression, scowling, and yawning categories, respectively, as shown in Figure 2. The mean of the confidence scores in each rated category suggested that the most difficult category to classify correctly by AI in the rated categories was blinking, whose confidence score was 0.51±0.35 (mean±SD).

Sample images of true and false classifications by AI are shown in Figure 3.

It took less than 0.1 s to classify an image. The L2 regularization value was 0 for the best performance.



**Figure 1:** The confusion matrix plot for the test dataset. The correct predictions were located in a diagonal matrix position.

## Discussion

Here, an AI classifier for deep learning with CNN of the original neural network architecture is developed using images of fetal faces. The accuracy for classifying fetal facial expressions was 0.984, which seemed excellent. The other statistical results, such as sensitivity, specificity, etc., also demonstrated high yields, as shown in Table 2. All statistical values were over 0.93 except for the scowling sensitivity value, 0.888, probably due to fewer cases. Although there were no significant differences for maternal age between the five categories, there were significant differences for gestational weeks between them as a bias because this was not a randomized study. This bias can be taken into account for interpreting this study.

The recognition of facial expressions by AI for adults has been investigated. A facial expression that can define human mental state and behavior and is used for security purposes and recognition of facial expressions by AI is used in domains such as healthcare, marketing,

**Table 2:** The statistic result for fetal facial expression by an original neural network for artificial intelligence (AI).

	Eye blinking	Mouthing	Face without any expression	Scowling	Yawning
The number of true positive	14	143	67	8	25
The number of true negative	246	116	190	252	236
The number of false positive	0	0	4	0	0
The number of false negative	1	2	0	1	0
Sensitivity	0.933	0.986	1.000	0.888	1.000
Specificity	1.000	1.000	0.979	1.000	1.000
Positive predictive value	1.000	1.000	0.943	1.000	1.000
Negative predictive value	0.995	0.983	1.000	0.996	1.000
Accuracy	0.996	0.992	0.985	0.996	1.000
F1 score	0.965	0.993	0.971	0.941	1.000
Informedness	0.933	0.986	0.979	0.889	1.000

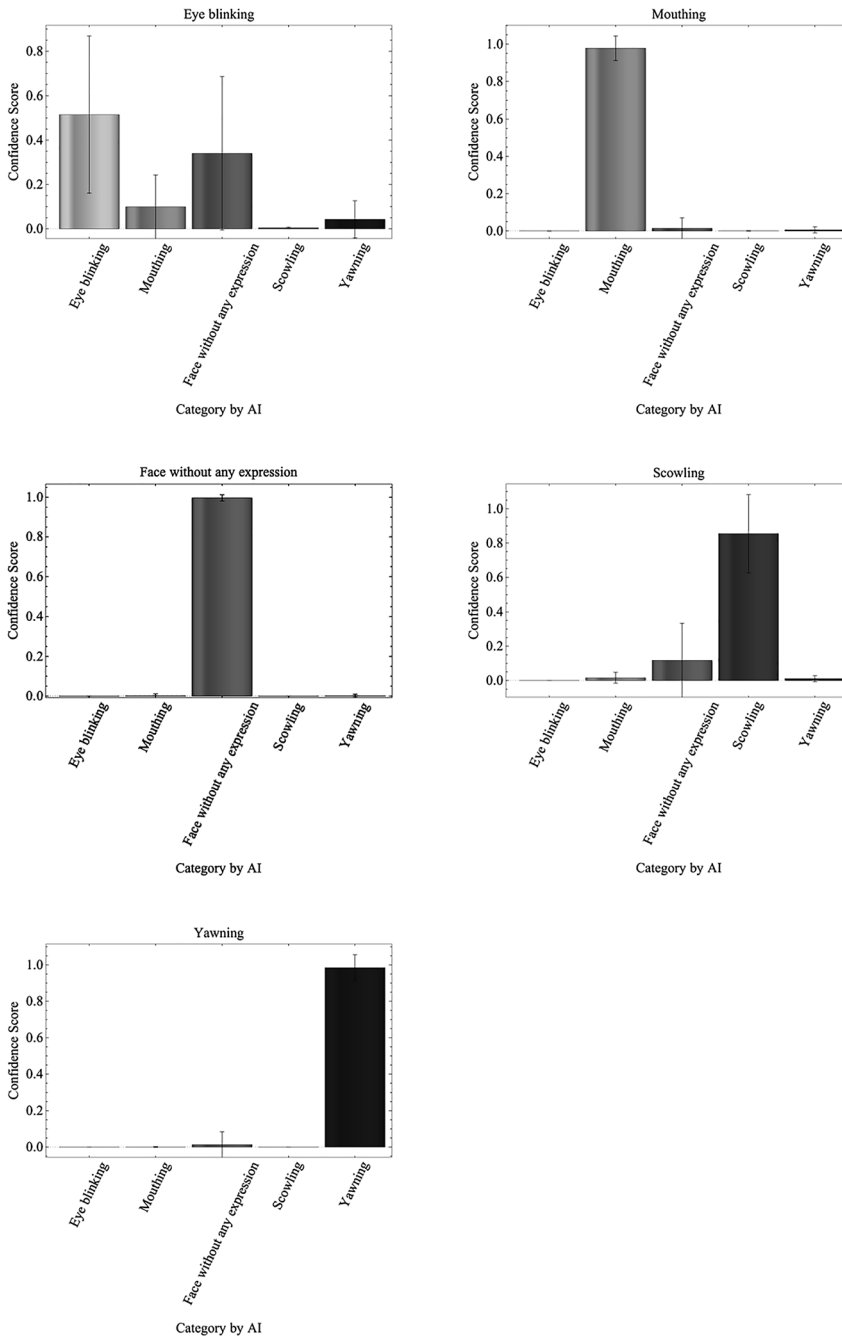
environment, safety, and social media [47]. Chen et al. reported that CNN as a deep learning architecture can extract the essential features of facial expression image and classified facial images into seven facial expressions [48]. According to reports, AI uses a variety of algorithms to reorganize adult facial expressions, and CNN seems to be mandatory [49]. Kim et al. reported that the accuracy of the AI facial expression recognition dataset was 0.965 [50].

The accuracy of fetal facial expression in this study was 0.985 which did not appear inferior compared to its result. However, there have been no reports regarding the recognition of fetal facial expressions by AI to our knowledge.

In this study, we define seven categories; eye blinking, mouthing, face without any expression, scowling, smiling, tongue expulsion, and yawning. AboEllail et al. [13] reported seven similar categories for fetal face expression that included sucking rather than the face without any expression, but sucking was rarely observed, and the face without any expression was often observed in our preliminary studies. There were only several cases of smiling and tongue expulsion in this study, and these two categories were removed before the creation of AI. They also reported other fetal facial expressions such as frown, sad, and funny face as emotion-like behavior. Therefore, the complete classification of fetal facial expression has not yet been established, probably because it could take time to observe the fetal face for a long time and the image classification could not be identical by doctors and sonographers. AI that is a computer program, however, has no bias in itself for classifying images. AI could demonstrate objective findings regarding fetal facial expressions, then research into the development of the fetal brain and central nervous system would be advanced.

The most difficult category to classify correctly by AI in the rated categories was eye blinking, whose confidence score was  $0.51 \pm 0.35$  (mean  $\pm$  SD), as shown in Figure 2. Eye blinking was sometimes classified as face without any expression, and the confidence score was  $0.35 \pm 0.34$  in the eye blinking category. The maximum values of the mean of the confidence scores of other categories were over 0.85. Trying to collect more eye blinking cases would improve the AI classifier in the future.

The limitations of this study should be considered. Firstly, the AI cannot correctly classify unknown images. Seven categories were used, but there could be other fetal facial expressions that could be meaningful in investigating fetal brain and central nervous system development. Those undiscovered and undefined categories and images could be needed to train AI for research and clinical practice in the future. Secondly, the feasibility of the recognition of fetal facial expression by AI in this study depends on the supervised data derived from Japanese fetuses. Since the anthropometric differences reflected in fetal facial expression can strongly affect AI creation, this AI might not be used directly for different anthropometric fetuses. However, I believe that the same algorithm could be available for other anthropometric fetuses. Thirdly, although the AI in this study demonstrated very good accuracy, there are still problems of incomplete datasets for seven categories because the smiling and tongue expulsion were rarely seen. These datasets will need to be collected. Fourthly, patients were not randomized in this pilot study, resulting in gestational weeks bias. Therefore, the AI recognition accuracies and frequencies of each fetal facial expression related to gestational weeks could not be analyzed. Randomization of patients across multiple facilities would be desirable in subsequent studies.



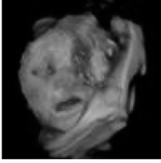








**Figure 2:** The profiles of the confidence scores for each category per image [mean±standard deviation (SD)]. The predicted category by artificial intelligence (AI) was correct in all categories. AI predicted blinking correctly with 0.51 of the score and neutral incorrectly with 0.34 of the score in the blinking category (left upper panel). The most difficult category to classify by AI in the qualified categories was blinking.

Since there is no gold-standard neural network architecture for CNN, AI itself has multimodality that can consolidate a series of disconnected, heterogeneous data. For example, regarding the diagnosis of AI-assisted colposcopy, the accuracy value by image only and by both image and human papillomavirus typing were 0.823 and 0.941, respectively [17, 18].

In other words, fetal facial expression can be classified not only by image but also by incorporated images with known parameters such as gestational age. This

could be an advantage of AI to recognize the facial expression of the fetus. When those advanced AIs are provided for the recognition of fetal facial expression using incorporated images with some parameters, investigations could be carried out for profiles of the relationship between fetal facial expressions and parameters such as parity, siblings, multiple pregnancies, maternal disease, maternal personality, maternal diet, fetal physical development, physical and mental development after birth, intelligence, score in school,

Category	True	False	AI classification
Eye blinking			Face without any expression
Mouthing			Face without any expression
Face without any expression			not available
Scowling			Face without any expression
Yawning			Face without any expression

**Figure 3:** The sample images of true and false classifications by artificial intelligence (AI). The AI classification for false cases in each category is shown. All neutral fetal facial expressions were correctly classified, then the false case in the neutral category is missing.

personality formation, etc. The findings could suggest to pregnant women how best to treat their fetus.

Since the fetal facial expression is believed to be essential for investigating fetal brain and central nervous system development non-invasively, AI can advance the investigation of fetal brain development using ultrasound.

**Research funding:** None declared.

**Author contributions:** All authors have accepted responsibility for the entire content of this manuscript and approved its submission. Miyagi Y and Hata T designed and coordinated the study; Miyagi Y and Hata T supervised the project; Hata T, Koyanagi Y and Bouno S acquired and validated data; Miyagi Y developed artificial intelligence software, analyzed and interpreted data, and wrote draft; Miyake T set up project administration; Miyagi Y, Hata T, Koyanagi A, Bouno S and Miyake T wrote the manuscript; and all authors approved the final version of the article.

**Competing interests:** Authors state no conflict of interest.

**Ethical approval:** This retrospective, non-interventional study was approved by the Institutional Review Board (IRB) of Miyake Clinic (IRB No. 2019-10).

## References

- Hata T, Dai SY, Marumo G. Ultrasound for evaluation of fetal neurobehavioural development: from 2-D to 4-D ultrasound. *Infant Child Dev* 2010;19:99–118.
- Hata T, Kanenishi K, Hanaoka U, Marumo G. HDlive and 4D ultrasound in the assessment of fetal facial expressions. *Donald Sch J Ultrasound Obstet Gynecol* 2015;9:44–50.
- Hata T. Current status of fetal neurodevelopmental assessment: 4D ultrasound study. *J Obstet Gynaecol Res* 2016;42:1211–21.
- Nijhuis JG. Fetal behavior. *Neurobiol Aging* 2003;24:S41–6.
- Prechtl HF. State of the art of a new functional assessment of the young nervous system: an early predictor of cerebral palsy. *Early Hum Dev* 1997;50:1–11.
- de Vries JIP, Visser GHA, Prechtl HFR. The emergence of fetal behaviour. I. Qualitative aspects. *Early Hum Dev* 1982;7:301–22.
- de Vries JIP, Visser GHA, Prechtl HFR. The emergence of fetal behaviour. II. Quantitative aspects. *Early Hum Dev* 1985;12:99–120.
- Prechtl HF. Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction. *Early Hum Dev* 1990;23:151–8.
- Prechtl HF, Einspieler C. Is neurological assessment of the fetus possible? *Eur J Obstet Gynecol Reprod Biol* 1997;75:81–4.
- Kuno A, Akiyama M, Yamashiro C, Tanaka H, Yanagihara T, Hata T. Three-dimensional sonographic assessment of fetal behavior in

- the early second trimester of pregnancy. *J Ultrasound Med* 2001; 20:1271–5.
11. Kurjak A, Miskovic B, Stanojevic M, Amiel-Tison C, Ahmed B, Azumendi G, et al. New scoring system for fetal neurobehavior assessed by three- and four-dimensional sonography. *J Perinat Med* 2008;36:73–81.
  12. Stanojevic M, Talic A, Miskovic B, Vasilj O, Shaddad AN, Ahmed B, et al. An attempt to standardize Kurjak's antenatal neurodevelopmental test: Osaka consensus statement. *Donald Sch J Ultrasound Obstet Gynecol* 2011;5:317–29.
  13. AboEllaail MAM, Hata T. Fetal face as important indicator of fetal brain function. *J Perinat Med* 2017;45:729–36.
  14. Hata T, Kanenishi K, AboEllaail MAM, Marumo G, Kurjak A. Fetal consciousness 4D ultrasound study. *Donald Sch J Ultrasound Obstet Gynecol* 2015;9:471–4.
  15. Reissland N, Francis B, Kumarendran K, Mason J. Ultrasound observations of subtle movements: a pilot study comparing fetuses of smoking and nonsmoking mothers. *Acta Paediatr* 2015;104:596–603.
  16. Miyagi Y, Miyake T. Potential of artificial intelligence for estimating Japanese fetal weights. *Acta Med Okayama* 2020;74: 483–93, PubMedID: 33361868.
  17. Miyagi Y, Takehara K, Nagayasu Y, Miyake T. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images combined with HPV types. *Oncol Lett* 2020;19:1602–10.
  18. Miyagi Y, Takehara K, Miyake T. Application of deep learning to the classification of uterine cervical squamous epithelial lesion from colposcopy images. *Mol Clin Oncol* 2019;11: 583–9.
  19. Miyagi Y, Habara T, Hirata R, Hayashi N. Predicting a live birth by artificial intelligence incorporating both the blastocyst image and conventional embryo evaluation parameters. *Artif Intell Med Imaging* 2020;1:94–107.
  20. Miyagi Y, Habara T, Hirata R, Hayashi N. Feasibility of artificial intelligence for predicting live birth without aneuploidy from a blastocyst image. *Reprod Med Biol* 2019;18:204–11.
  21. Miyagi Y, Habara T, Hirata R, Hayashi N. Feasibility of deep learning for predicting live birth from a blastocyst image in patients classified by age. *Reprod Med Biol* 2019;18:190–203.
  22. Miyagi Y, Habara T, Hirata R, Hayashi N. Feasibility of predicting live birth by combining conventional embryo evaluation with artificial intelligence applied to a blastocyst image in patients classified by age. *Reprod Med Biol* 2019;18:344–56.
  23. Miyagi Y, Fujiwara K, Oda T, Miyake T, Coleman RL. Studies on development of new method for the prediction of clinical trial results using compressive sensing of artificial intelligence. In: Ferreira MAM, editor. *Theory and practice of mathematics and computer science*. Hooghly, West Bengal, India: Book Publisher International; 2020. pp. 101–8. <https://doi.org/10.9734/bpi/tpmcs/v2>.
  24. Miyagi Y, Fujiwara K, Oda T, Miyake T, Coleman RL. Development of new method for the prediction of clinical trial results using compressive sensing of artificial intelligence. *J Biostat Biometric Appl* 2018;3:202.
  25. Miyagi Y, Tada K, Yasuhi I, Maekawa Y, Okura N, Kawakami K, et al. New method for determining fibrinogen and FDP threshold criteria by artificial intelligence in cases of massive hemorrhage during delivery. *J Obstet Gynaecol Res* 2020;46: 256–65.
  26. Bengio Y, Courville A, Vincent P. Representation learning: a review and new perspectives. *IEEE Trans Pattern Anal Mach Intell* 2013;35:1798–828.
  27. LeCun YA, Bottou L, Orr GB, Müller KR. Efficient backprop. In: Montavon G, Orr GB, Müller KR, editors. *Neural networks: tricks of the trade*. Heidelberg, Berlin: Springer; 2012. pp. 9–48. [https://doi.org/10.1007/978-3-642-35289-8\\_3](https://doi.org/10.1007/978-3-642-35289-8_3).
  28. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE* 1998;86:2278–324.
  29. LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation applied to handwritten zip code recognition. *Neural Comput* 1989;1:541–51.
  30. Serre T, Wolf L, Bileschi S, Riesenhuber M, Poggio T. Robust object recognition with cortex-like mechanisms. *IEEE Trans Pattern Anal Mach Intell* 2007;29:411–26.
  31. Wiatowski T, Bölcskei H. A mathematical theory of deep convolutional neural networks for feature extraction. *IEEE Trans Inf Theor* 2017;64:1845–66.
  32. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 2014;15:1929–58.
  33. Nowlan SJ, Hinton GE. Simplifying neural networks by soft weight-sharing. *Neural Comput* 1992;4:473–93.
  34. Ciresan DC, Meier U, Masci J, Gambardella LM, Schmidhuber J. Flexible, high performance convolutional neural networks for image classification. In: *Proceedings of the twenty-second international joint conference on artificial intelligence*, Barcelona, Spain; 2011. p. 1237–42.
  35. Scherer D, Müller A, Behnke S. Evaluation of pooling operations in convolutional architectures for object recognition. In: Diamantaras K, Duch W, Iliadis LS, editors. *Artificial neural networks – ICANN 2010*. Lecture notes in computer science. Berlin, Heidelberg: Springer; 2010. pp. 92–101. [https://doi.org/10.1007/978-3-642-15825-4\\_10](https://doi.org/10.1007/978-3-642-15825-4_10).
  36. Huang FJ, LeCun Y. Large-scale learning with SVM and convolutional for generic object categorization. In: *IEEE Computer Society conference on computer vision and pattern recognition*. IEEE, New York, USA; 2006. pp. 284–91. <https://doi.org/10.1109/CVPR.2006.164>.
  37. Zheng Y, Liu Q, Chen E, Ge Y, Zhao JL. Time series classification using multi-channels deep convolutional neural networks. In: Li F, Li G, Hwang S, Yao B, Zhang Z, editors. *Web-age information management*. WAIM 2014. Lecture notes in computer science. Cham: Springer; 2014. pp. 298–310. [https://doi.org/10.1007/978-3-319-08010-9\\_33](https://doi.org/10.1007/978-3-319-08010-9_33).
  38. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *Nature* 2015;518:529–33.
  39. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, Boston, USA, 2015. Computer Vision Foundation; 2015. pp. 1–9.
  40. Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: *Proceedings of the fourteenth international conference on artificial intelligence and statistics (AISTATS)*, Lauderdale, USA, 2011. AISTATS; 2011. pp. 315–23.
  41. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)*, Haifa, Israel. Omni Press; 2010. pp. 807–14.

42. Ioff S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. <https://arxiv.org/abs/1502.03167v3>.
43. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Proceedings of the 25th international conference on neural information processing systems; 2012. pp. 1097–105. <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
44. Bridle JS. Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. In: Soulié FF, Héroult J, editors. Neurocomputing. Berlin, Heidelberg: Springer; 1990. pp. 227–36. [https://doi.org/10.1007/978-3-642-76153-9\\_28](https://doi.org/10.1007/978-3-642-76153-9_28).
45. Powers DMW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness & correlation. *J Mach Learn Technol* 2011;2:37–63.
46. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3: 32–5.
47. Dixit AN, Kasbe T. A survey on facial expression recognition using machine learning techniques. In: 2nd international conference on data, engineering and applications (IDEA); 2020. pp. 1–6. <https://doi.org/10.1109/IDEA49133.2020.9170706>.
48. Chen X, Yang X, Wang M, Zou J. Convolution neural network for automatic facial expression recognition. In: 2017 International conference on applied system innovation (ICASI); 2017. pp. 814–17. <https://doi.org/10.1109/ICASI.2017.7988558>.
49. Hu M, Yang C, Zheng Y, Wang X, He L, Ren F. Facial expression recognition based on fusion features of center-symmetric local signal magnitude pattern. *IEEE Access* 2019;7:118435–45.
50. Kim J, Kim B, Roy PP, Jeong DM. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. *IEEE Access* 2019;7:41273–85.