A Computer-Aided Diagnosis System of Fetal Nucleated Red Blood Cell With Convolutional Neural Network

Chao Sun
Department of Obstetrics and Gynecology
Ruijie Wang
School of Electronic and Information Engineering
Lanbo Zhao
Department of Obstetrics and Gynecology
Lu Han
Department of Obstetrics and Gynecology
Sijia Ma
Department of Obstetrics and Gynecology
Dongxin Liang
Department of Obstetrics and Gynecology
Lei Wang
Department of Obstetrics and Gynecology
Xiaoqian Tuo
Department of Obstetrics and Gynecology
Yu Zhang
Department of Obstetrics and Gynecology
Dexing Zhong
School of Electronic and Information Engineering
Qiling Li (liqiling@mail.xjtu.edu.cn)
Department of Obstetrics and Gynecology https://orcid.org/0000-0003-3763-4456

Research

Keywords: fetal nucleated red blood cells, computer-aided diagnosis system, region of interest, convolutional neural network

Posted Date: January 29th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-154341/v1

Version of Record: A version of this preprint was published at Archives of Pathology & Laboratory Medicine on March 16th, 2022. See the published version at https://doi.org/10.5858/arpa.2021-0142-OA.

Article

A Computer-Aided Diagnosis System of Fetal Nucleated Red Blood Cell With Convolutional Neural Network

Chao Sun[1], Ruijie Wang[2], Lanbo Zhao[1], Lu Han[1], Sijia Ma[1], Dongxin Liang[1], Lei Wang[1], Xiaoqian Tuo[1],Yu Zhang[1], Dexing Zhong[2,3,4]* & Qiling Li[1]*

[1]Department of Obstetrics and Gynecology, The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an 710061, China. [2]School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China. [3]State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, [4]China. Pazhou Lab, Guangzhou 510335, China*Email: liqiling@mail.xjtu.edu.cn; bell@xjtu.edu.cn

Abstract

Background: The rapid recognition of fetal nucleated red blood cells (fNRBCs) present considerable challenges.

Objective: To establish a computer-aided diagnosis system (CAD) for rapid recognition of fNRBCs by a convolutional neural network (CNN).

Methods: We adopted density gradient centrifugation and magnetic-activated cell sorting to extract fNRBCs from umbilical cord blood samples. A cell-block method was used to embed fNRBCs for routine formalin-fixed paraffin sectioning and hematoxylin-eosin stains. Then we proposed a CAD-based on CNN to automatically learn discriminative features and recognize fNRBCs. Region of interest [1] extraction methods were used to automatically segment individual cells in cell slices. The discriminant information from ROIs was encoded into a feature vector. The prediction network provided a pathological diagnosis.

Results: Totally, 4760 pictures of fNRBCs from 260 cell-slides of 4 umbilical cord blood samples were collected. On the premise of 100% accuracy in the training set (3720 pictures) the sensitivity, specificity, and accuracy of cellular intelligent recognition were 96.5%, 100%, and 98.5% in the test set (1040 pictures).

Conclusion: We present a CAD system for effective and accurate fNRBCs recognition based on CNN.

Keywords: fetal nucleated red blood cells; computer-aided diagnosis system; region of interest; convolutional neural network

Introduction

The clinical application of fetal nucleated red blood cells (fNRBCs) during pregnancy is classified into two main categories[2,3]. One is the evaluation of chronic tissue hypoxia in pregnant women by counting fNRBCs in the umbilical and peripheral blood of newborns. Because chronic fetal tissue hypoxia results in increased levels of erythropoietin, which, in turn, leads to the stimulation of erythropoiesis and increases numbers of fetal circulating fNRBCs[2,4-6]. The chronic tissue hypoxia is usually seen in intrauterine growth restriction, maternal hypertension, preeclampsia, maternal smoking, Rh isoimmunization, and maternal diabetes[6-10]. Another is to screen and extract fNRBCs from maternal peripheral blood for non-invasive prenatal diagnosis (NIPD)[11-13]. Because fNRBCs possess cell-surface markers for the enrichment and bear a complete complement of fetal nuclear genes[14].

In the past 2 decades, fNRBCs have been identified from the umbilical cord and maternal peripheral blood with the use of various enrichment techniques, such as density gradient centrifugation (DGC) [15,16], fluorescence-activated cell sorting (FACS)[17], magnetic-activated cell sorting (MACS)[18], dielectrophoresis, and microfluidics-based technology[19,20]. However, complex procedures limit extensive studies. Despite considerable progress, reproducibility, and reliability of isolation and detection of fetal cells from maternal blood remain poor. That was attributed to the rarity and variability of fetal cells among pregnancies[21,22]. Due to the lack of a specific screening method, the a few fNRBCs recognition in vast negative cells remains difficult and represents a huge manual burden on researchers.

In recent years, the development of computer-aided diagnosis(CAD) and medical image processing has resulted in the emergence of the field of computational

pathology[23]. The difficulty of fNRBCs recognition may benefit greatly from urgent digital revolution[24]. Based on the combination of deep learning (DL) and multi-medical specialties, CAD has rapidly gained popularity and led to substantial progress in fields such as radiology, ophthalmology, and pathology[25-30]. DL-based algorithms have demonstrated remarkable progress in image recognition tasks. As the most prevalent type of DL structure, the convolution neural network (CNN) has a natural advantage in utilizing the 2D structure of an image[31]. CNN-based CAD systems have been reported to surpass human performance[32], are widely used for analysis in medical image[33]. Inspired by this, we investigate whether a CNN-based CAD system can be employed as a classifier to automatically recognize fNRBCs in the slice of umbilical cord cells.

Materials and Methods

Ethics statement. This study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University(XJTU1AF-CRS-2015-001). Related informed consent was obtained from the patients before the study, and all the protocols complied with the ethical principles for research that involves human subjects of the Helsinki Declaration for medical research[34].

Cord blood samples (Figure 1A-a). All umbilical cord blood samples were collected from normal term deliveries (≥37 weeks). Approximately 9 ml of cord blood was collected into anti-coagulant K2-EDTA tubes (BD Vacutainer 366643) containing a proprietary preservative.

fNRBCs enrichment (Figure 1A-b). Umbilical cord blood samples were processed within 2 hours of collection and mononuclear cells were isolated by DGC (centrifuged at 1,500 rpm for 30 min) with Histopaque-1077 (Sigma Chemical, St. Louis, MO, USA)[35]. fNRBCs were magnetically labeled with an anti-CD71 monoclonal antibody (Miltenyi Biotec, Germany), and positively selected by MACS (Number: 130-091-632, Miltenyi, Biotec, Germany) according to the protocoll provided by the manufacturer[36].

fNRBCs fixation by a cell-block technique (Figure1A-c) and hematoxylin-eosin (HE) staining. The fNRBCs samples were washed twice with PBS, resuspended in 2 mL of phosphate buffer saline, and were centrifuged at 2000 rpm at room temperature for 10 min. Then, the temperature was increased to 40 °C, and the cell-rich layer was collected. The samples were transferred to the bottom with the diluted solution (Xi'an Meijiajia, China), loaded into the Li-Shi Thin Prep Liquid-based Cytology and Tissue

Embedding Machine (Xi'an Meijiajia, China), and removed after centrifugation at 2000 rpm for 10 min. After standing at room temperature for 10 min, the samples were taken out. The parts without the cell-rich layer were cut off, and the left parts were stored in the embedding box. According to routine protocols[37], HE staining were performed on 5-μm sections of formalin-fixed, paraffin-embedded tissues.

Image acquisition and processing (Figure 1A-d). After HE staining, the Pathological Section Scanner (Pannoramic DESK, P-MIDI, P250, P1000; 3D HISTECH； Hungary) scanned the sections. The main unique cytological characteristics as follows: The diameter of fNRBCs（9-12 μm）was between small lymphocytes (5-8 μm) and neutrophils（10-20 μm）[38]. The ratio of the nucleus to the cytoplasm was less than 1/2. The nucleus was dense, massive[39], and off-center. The cytoplasm was orthochromatic nongranular[40]. Two senior pathologists independently identified the cells. Only when both pathologists judged a cell as fNRBC at the same time, the cell was recorded as positive. Otherwise, the cell was negative.

Cellular-level region of Interest (ROI) extraction. A semi-automated ROI extraction algorithm based on global threshold segmentation and watershed algorithm was proposed[41]. First, a Gaussian low-pass filter was applied for image pre-processing. Our goal was to collect single-cell regions as ROI data sets. To reduce the complexity of ROI extraction, adaptive thresholding methods, and mathematical morphology operations were adopted to segment fNRBCs. The adaptive threshold T was calculated by the following formula:

$$T = \frac{Iocs(x) + (256 - Iocs(x))}{3.06}$$

where, as Figure 2A showed the first extreme point on the histogram of the grayscale distribution was denoted as *locs*, and *locs(x)* represented the corresponding abscissa.

Considering the information in the grayscale image (Figure 2A) , an improved watershed method based on adaptive thresholding was proposed. First, information on the image gradient was used as prior knowledge, and the watershed algorithm was rendered sensitive to the small extreme line response[42]. Then, the mathematical morphology technique was used to remove cell debris, and over-segmentation was eliminated by bottleneck detection as Figure 1B-a.

Data augmentation. The data augmentation technique usually adopted by adding noise or applying geometric transformations to existing pictures. In this research, we performed a rotation by 90°, 180°, and 270°clockwise. Besides, we randomly flipped the images horizontally. Respectively, thereby we expanded the original data set by 8 times (Figure 1B-b).

Prediction network. We encoded the ROI set as feature vectors, and CNN was used to detect cell-level features and perform detection classification. We proposed a predictive network to perform the classification task of fNRBC images.

The predictive network (P-net) was an end-to-end trainable network[43], which consisted of three convolutional layer blocks and three maximum pooling layers. P-net could automatically learn the feature representation of the image, and finally, the fully connected neural network applies these meaningful representations to diagnosis. Due to the limitations of linear expression, many features of the original input were not preserved. We combined the data of the input image to generate more features of the image, which conferred greater stability and efficiency to the network.

We chose the rectified linear unit (ReLU) function as the activation function.

Since the size of the ROI patch was different, we filled the pixels around the patch until the size became 120×120, and input them to the network. Through the training of 10,000 samples, P-net was fine-tuned on a domain-specific dataset. The prediction network framework was shown in Figure 1C.

Statistics. We used accuracy, sensitivity and specificity to evaluate the results. A series of parameters were calculated as true positives (TP, correctly classified as positive), true negatives (TN, correctly classified as negative), false positives (FP, error classified as positive), and false negatives (FN, error classified as negative), the standard was defined as follows:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

$$\text{Sensitivity} = \frac{TP}{(TP + FN)}$$

$$\text{Specificity} = \frac{TN}{(TN + FP)}$$

We chose the precision-recall curve (Figure 2B) to evaluate the efficiency of P-net and traditional CNN networks (Supplementary Figure 1). It showed the performance of different networks between precision and recall. The formula for precision was:

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

And recall was related to the true positive rate and the same as the definition of

sensitivity. The larger the area under the curve, the higher the recall and precision was represented. Here, Net1 represented a traditional CNN network, while Net2 referred to the network proposed.

Results

Baseline characteristics. This study was conducted at the First Affiliated Hospital of Xi'an Jiaotong University. The study included 4 pregnant women, who all delivered a single mature neonates. The mean maternal age was 28.75 years, the mean gestational age was 38 weeks and 6 days, and the mean birth weight was 3447.5 gram. Table 1 listed the demographic data of these participants.

Cell-block and HE staining of fNRBCs. Eight cell-blocks of fNRBCs from 4 umbilical cord blood samples were made. Thirty-two HE stained cell slices were obtained and scanned images using Pathological Section Scanner (Leica SCN 400, Germany). Under the microscope, the sliced background was clean and free of impurities, the cells were evenly distributed without overlap, and the cell staining was clear enough for recognition (Figure 2C). fNRBCs were observed in all umbilical cord blood samples. Most fNRBCs showed red cytoplasm and off-center nucleus, which were different in the histologic features of lymphocytes and neutrophils (Figure 2C).

Data set. With average slide dimensions of 1,651 × 1,209 pixels (height × width), 260 sample images were acquired. Initially, 595 pictures of fNRBCs were collected. The training set and test set were randomly split in a 7:3 ratio. 465 and 130 pictures were used for training and testing. After data augmentation, we obtained 4760 pictures of fNRBCs. Finally, training and testing pictures reached 3,720 and 1,040.

Evaluation metrics. We validated the CNN model on this dataset. The total number of positive and negative samples was similar. When the output of the neural network was close to 1, the sample was likely to be positive. Conversely, an output value close

to 0 indicated that the sample belonged to the negative group (Figure 1C).

Table 2 showed the comparison of the results of the method used in the experimental process but eventually eliminated and the existing method. The experiments proved that data enhancement effectively improve the classification effect of the P-net. Finally, on the premise of 100% accuracy in the training set, the test set was observed to attain 96.5% sensitivity, 100% specificity, and 98.5% accuracy (Table 2 Scheme variables part). Without data augmentation, the test set was observed to attain 90% accuracy, 91% sensitivity, and 89% specificity. After data augmentation, the accuracy, sensitivity, and specificity were prompted to 98.5%, 96.5%, and 98.5% (Table 2 Methods part).

Discussion

To the best of our knowledge, this is the first time that a CNN-based CAD system was applied for automatic recognition of fNRBCs. As a type of DL network that evolved from the multilayer perceptron, CNN can be effectively applied in areas such as image recognition and classification, object detection, and natural language processing[31]. Compared with the traditional multilayer perceptron, which usually used a fully connected network; CNN performed better in the field of image processing due to its structural characteristics of local connectivity, shared weights, and downsampling[30]. Our results indicated that the proposed CNN model was able to quickly and easily identify fNRBCs. The CAD system proposed obtained 98.5% accuracy and 96.5% sensitivity, which provides a new idea for the rapid identification of cells.

As associated with an inevitably high loss of FNRBCs, between 30% and 70%[44], DGC is usually used to deplete the overwhelming abundance of maternal red blood cells (density of fNRBCs is similar to that of mononuclear cells and slightly lower than red blood cells) as the initial enrichment step[45]. As a simple, fast, and economical bench-top technique with the capability to process large cell numbers, MACS is usually selected as a further enrichment method[46]. Transferrin receptor (CD71) is a common positive marker[47,48], which was expressed on the entire erythroid lineage, activated lymphocytes, monocytes, trophoblasts, any cell incorporating iron and definitive fNRBCs in maternal and cord blood. In our study, DGC and MACS (anti-CD71) method were used to separate f{Choolani, 2003 #119}{Ganshirt, 1998 #69}NRBCs from maternal cord blood. Totally, 4760 pictures of fNRBCs from 260 cell-slides of 4 umbilical cord blood samples were collected.

Moreover, we report a effective method for the long-term preservation of fNRBCs. Cell-block preparations have been used as a complementary technique for increasing diagnostic accuracy in many fields[49], such as endometrial cytology, malignant pleural effusion, and needle aspiration cytology of thyroid gland[50,51]. We first proposed a cell-block technique for fixation of fNRBC samples. This technique could ensure a uniform distribution of the enriched fNRBCs in the wax block, which is convenient for the identification and isolation of individual fetal cells at a later stage. Also, the cell slices generated by this technique have no background interference to subsequent immunohistochemistry, fluorescent in situ hybridization, and other molecular pathology assays. Our method (Cell-Block technique) can not only preserve fNRBCs for a long time but also facilitate repeated tests using the same sample.

Fixed threshold image segmentation was not suitable for all images, for the image would be uneven and may fade over time. To reduce the complexity in the image classification algorithm, we proposed an adaptive ROI extraction method for fNRBC images. Since the global threshold algorithm could not distinguish adjacent cells, we chose the watershed algorithm to detect the single cell. Besides, we have comprehensively utilized the visual information perceived by the network and constructed a novel pathological recognition network, which would have significant contributions in improving the means and methods of non-invasive medical diagnostics.

Actually, in some medical domains, large-scale labeled datasets are unavailable. However, the excellent performance of a deep neural network greatly relies on the scale of the labeled dataset[52-54]. Therefore, some techniques have been performed to prevent overfitting when training a neural network on a small dataset, such as data

augmentation. To perform data augmentation, the simplest way is to add noise or apply a geometric transformation to existing data. Applying noise and transformations to images of lesions makes sense since this kind of data is very likely to be affected by all sorts of noise and can be found in different sizes and orientations. Geometric transformation is based on image manipulation processes, including flipping, cropping, rotation, color space transformation, and so on. Lai, L et al.[31] and Chougrad, H et al.[52] performed geometric transformation for data augmentation. These transformations boost the models to learn better[53]. In our study, the accuracy, sensitivity, and specificity of the model were prompted through data augmentation significantly.

The present study had two limitations. Due to the rare of fNRBCs in maternal peripheral blood, it is not enough for the initial stage of the AI system establishment. Therefore, the umbilical cord blood of pregnant women was selected as the sample for both input and verification. Besides, our study was a preliminary attempt to apply DL to fNRBCs recognition. An automatic and application-oriented CAD system remains to be developed in further work.

Conclusion

In the aspect of NIPD, fNRBCs are attracting more and more attention. Based on traditional extraction methods, we introduce the cell-block technology, which makes the long-term preservation of samples and repeated experiments possible. Besides, the establishment of the CAD system provides a possibility for fast recognition of fNRBCs. In the future, we would conduct further investigations on maternal peripheral blood, and continue to optimize the system, such that it can be devoted to fNRBC recognition in NIPD.

List of abbreviations

fNRBC: fetal nucleated red blood cell

NIPD: non-invasive prenatal diagnosis

DGC: density gradient centrifugation

FACS: fluorescence-activated cell sorting

MACS: magnetic-activated cell sorting

CAD: computer-aided diagnosis

DL: deep learning

CNN: convolution neural network

HE: hematoxylin-eosin

ROI: region of interest

P-net: predictive network

ReLU: rectified linear unit

TP: true positives

TN: true negatives

FP: false positives

FN : false negatives

CW: clockwise

Ethics and consent

This study was approved by the Ethics Committee of the First Affiliated Hospital of Xi'an Jiaotong University(XJTU1AF-CRS-2015-001). Related informed consent was obtained from the patients before the study.

Consent for publication

There are not details, images, or videos relating to an individual person.

Competing interests

The authors declare no competing interests.Correspondence and requests for materials should be addressed to Q.L. or D.Z.

Funding

This work was supported by the Clinical Research Award of the First Affiliated Hospital of Xi'an Jiaotong University, China (XJTU1AF-2018-017, XJTU1AF-CRF-2019-002), the Major Basic Research Project of Natural Science of Shaanxi Provincial Science and Technology Department (2017ZDJC-11), the Key Research and Development Project of Shaanxi Provincial Science and Technology Department (2017ZDXM-SF-068, 2019QYPY-138), and Shaanxi Provincial Collaborative Technology Innovation Project (2017XT-026, 2018XT-002), the Basic Research and Development Project of Shaanxi Provincial Science and Technology Department (2020JQ-530). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Q.L. and D.Z. conceived and designed the study. C.S., L.Z., L.H., Y.Z.and S.M. performed the laboratory experiments. L.W. finished image acquisition and processing. R.W. analyzes and interprets data. C.S. and R.W. wrote the first draft of the manuscript. D.L. and X.T. revised the manuscript.

Acknowledgements

My deepest gratitude goes first and foremost to the 4 participants for providing specimens for our scientific research. Then my thanks would go to partners for their work on the experiment.

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1 Hennerbichler, S. *et al.* Fetal nucleated red blood cells in peripheral blood of pregnant women: detection and determination of location on a slide using laser-scanning cytometry. *Prenat Diagn* **23**, 710-715, doi:10.1002/pd.668 (2003).

2 Kil, T. H. *et al.* A study on the measurement of the nucleated red blood cell (nRBC) count based on birth weight and its correlation with perinatal prognosis in infants with very low birth weights. *Korean J Pediatr* **54**, 69-78, doi:10.3345/kjp.2011.54.2.69 (2011).

3 Krajewski, P., Welfel, E., Kalinka, J., Pokrzywnicka, M. & Kwiatkowska, M. [Evaluation of the relationship between circulating nucleated red blood cells count and inborn infection in neonates]. *Ginekol Pol* **79**, 17-22 (2008).

4 Li, J. *et al.* Nucleated red blood cell counts: an early predictor of brain injury and 2-year outcome in neonates with hypoxic-ischemic encephalopathy in the era of cooling-based treatment. *Brain Dev* **36**, 472-478, doi:10.1016/j.braindev.2013.06.012 (2014).

5 Masoudi, Z., Akbarzadeh, M., Vaziri, F., Zare, N. & Ramzi, M. The effects of decreasing maternal anxiety on fetal oxygenation and nucleated red blood cells count in the cord blood. *Iran J Pediatr* **24**, 285-292 (2014).

6 Boskabadi, H. *et al.* Nucleated red blood cells count as a prognostic biomarker in predicting the complications of asphyxia in neonates. *J Matern Fetal Neonatal Med* **30**, 2551-2556, doi:10.1080/14767058.2016.1256988 (2017).

7 Constantino, B. T. & Rivera, G. K. Q. Cutoff Value for Correcting White Blood Cell Count for Nucleated Red Blood Cells: What is it? Why is it Important? *Lab Med* **50**, e82-e90, doi:10.1093/labmed/lmz016 (2019).

8 Davari-Tanha, F., Kaveh, M., Nemati, S., Javadian, P. & Salmanian, B. Nucleated red blood cells count in pregnancies with idiopathic intra-uterine growth restriction. *J Family Reprod Health* **8**, 77-81 (2014).

9 Walsh, B. H., Boylan, G. B., Dempsey, E. M. & Murray, D. M. Association of nucleated red blood cells and severity of encephalopathy in normothermic and hypothermic infants. *Acta Paediatr* **102**, e64-67, doi:10.1111/apa.12086 (2013).

10 Gasparovic, V. E., Ahmetasevic, S. G. & Colic, A. Nucleated red blood cells count as first prognostic marker for adverse neonatal outcome in severe preeclamptic pregnancies. *Coll Antropol* **36**, 853-857 (2012).

11 Breman, A. M. *et al.* Evidence for feasibility of fetal trophoblastic cell-based noninvasive prenatal testing. *Prenat Diagn* **36**, 1009-1019, doi:10.1002/pd.4924 (2016).

12 Wei, X. *et al.* Highly sensitive and rapid isolation of fetal nucleated red blood cells with microbead-based selective sedimentation for non-invasive prenatal diagnostics. *Nanotechnology* **29**, 434001, doi:10.1088/1361-6528/aad8c4 (2018).

13 Feng, C. *et al.* Non-invasive Prenatal Diagnosis of Chromosomal Aneuploidies and Microdeletion Syndrome Using Fetal Nucleated Red Blood Cells Isolated by Nanostructure Microchips. *Theranostics* **8**, 1301-1311, doi:10.7150/thno.21979 (2018).

14 Troeger, C., Holzgreve, W. & Hahn, S. A comparison of different density gradients and antibodies for enrichment of fetal erythroblasts by MACS. *Prenat Diagn* **19**, 521-526 (1999).

15 Mavrou, A. *et al.* Identification of nucleated red blood cells in maternal circulation: a second step in screening for fetal aneuploidies and pregnancy complications. *Prenat Diagn* **27**, 150-153, doi:10.1002/pd.1640 (2007).

16 Kovalak, E. E., Dede, F. S., Gelisen, O., Dede, H. & Haberal, A. Nonreassuring fetal heart rate patterns and nucleated red blood cells in term neonates. *Arch Gynecol Obstet* **283**, 1005-1009, doi:10.1007/s00404-010-1517-y (2011).

17 Sohda, S. *et al.* The proportion of fetal nucleated red blood cells in maternal blood: estimation by FACS analysis. *Prenat Diagn* **17**, 743-752 (1997).

18 Ganshirt, D. *et al.* Enrichment of fetal nucleated red blood cells from the maternal circulation for prenatal diagnosis: experiences with triple density gradient and MACS based on more than 600 cases. *Fetal Diagn Ther* **13**, 276-286, doi:10.1159/000020854 (1998).

19 Zhang, H. *et al.* Frequency-enhanced transferrin receptor antibody-labelled microfluidic chip (FETAL-Chip) enables efficient enrichment of circulating nucleated red blood cells for non-invasive prenatal diagnosis. *Lab Chip* **18**, 2749-2756, doi:10.1039/c8lc00650d (2018).

20 Ma, G. C. *et al.* A Silicon-based Coral-like Nanostructured Microfluidics to Isolate Rare Cells in Human Circulation: Validation by SK-BR-3 Cancer Cell Line and Its Utility in Circulating Fetal Nucleated Red Blood Cells. *Micromachines (Basel)* **10**, doi:10.3390/mi10020132 (2019).
21 Xiaoyan, X. & Hanping, C. Fetal nucleated red blood cells in maternal peripheral blood and gestational age. *Int J Gynaecol Obstet* **87**, 143-144, doi:10.1016/j.ijgo.2004.07.018 (2004).
22 Kuo, P. L. Frequencies of fetal nucleated red blood cells in maternal blood during different stages of gestation. *Fetal Diagn Ther* **13**, 375-379, doi:10.1159/000020873 (1998).
23 Fuchs, T. J., Wild, P. J., Moch, H. & Buhmann, J. M. Computational pathology analysis of tissue microarrays predicts survival of renal clear cell carcinoma patients. *Med Image Comput Comput Assist Interv* **11**, 1-8, doi:10.1007/978-3-540-85990-1_1 (2008).
24 Hajdu, S. I. A note from history: microscopic contributions of pioneer pathologists. *Ann Clin Lab Sci* **41**, 201-206 (2011).
25 Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. & Aerts, H. Artificial intelligence in radiology. *Nat Rev Cancer* **18**, 500-510, doi:10.1038/s41568-018-0016-5 (2018).
26 Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B. S., Waldstein, S. M. & Bogunovic, H. Artificial intelligence in retina. *Prog Retin Eye Res* **67**, 1-29, doi:10.1016/j.preteyeres.2018.07.004 (2018).
27 Zhou, L. Q. *et al.* Lymph Node Metastasis Prediction from Primary Breast Cancer US Images Using Deep Learning. *Radiology*, 190372, doi:10.1148/radiol.2019190372 (2019).
28 Merone, M., Sansone, C. & Soda, P. A computer-aided diagnosis system for HEp-2 fluorescence intensity classification. *Artif Intell Med* **97**, 71-78, doi:10.1016/j.artmed.2018.11.002 (2019).
29 Ha, T. *et al.* Comparison of the diagnostic performance of abbreviated MRI and full diagnostic MRI using a computer-aided diagnosis (CAD) system in patients with a personal history of breast cancer: the effect of CAD-generated kinetic features on reader performance. *Clin Radiol* **74**, 817 e815-817 e821, doi:10.1016/j.crad.2019.06.025 (2019).
30 Tanaka, H., Chiu, S. W., Watanabe, T., Kaoku, S. & Yamaguchi, T. Computer-aided diagnosis system for breast ultrasound images using deep learning. *Phys Med Biol* **64**, 235013, doi:10.1088/1361-6560/ab5093 (2019).
31 Lai, L., Cai, S., Huang, L., Zhou, H. & Xie, L. Computer-aided diagnosis of pectus excavatum using CT images and deep learning methods. *Sci Rep* **10**, 20294, doi:10.1038/s41598-020-77361-y (2020).
32 He, K., Zhang, X., Ren, S. & Jian, S. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*
33 Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Med Image Anal* **42**, 60-88, doi:10.1016/j.media.2017.07.005 (2017).
34 Issue Information-Declaration of Helsinki. *J Bone Miner Res* **33**, BM i-BM ii, doi:10.1002/jbmr.3259 (2018).
35 Wang, J. Y. *et al.* Fetal nucleated erythrocyte recovery: fluorescence activated cell sorting-based positive selection using anti-gamma globin versus magnetic activated cell sorting using anti-CD45 depletion and anti-gamma globin positive selection. *Cytometry* **39**, 224-230 (2000).
36 Yang, Y. H. *et al.* Prenatal genetic diagnosis from maternal blood: simultaneous immunophenotyping and FISH of fetal nucleated erythrocytes isolated by negative and positive magnetic activated cell sorting. *Yonsei Med J* **41**, 258-265, doi:10.3349/ymj.2000.41.2.258 (2000).
37 Serafini, S. *et al.* Is hematoxylin-eosin staining in rectal mucosal and submucosal biopsies still useful for the diagnosis of Hirschsprung disease? *Diagn Pathol* **12**, 84, doi:10.1186/s13000-017-0673-9 (2017).
38 Mohamed, H., Turner, J. N. & Caggana, M. Biochip for separating fetal cells from maternal circulation. *J Chromatogr A* **1162**, 187-192, doi:10.1016/j.chroma.2007.06.025 (2007).
39 Huang, R. *et al.* A microfluidics approach for the isolation of nucleated red blood cells (NRBCs) from the peripheral blood of pregnant women. *Prenat Diagn* **28**, 892-899, doi:10.1002/pd.2079 (2008).
40 Zou, L., Ye, X., Xu, K. & Zhu, J. Isolation of fetal nucleated red blood cells from maternal blood. *J Tongji Med Univ* **20**, 169-171, doi:10.1007/bf02887064 (2000).
41 Khan, A. U. M., Torelli, A., Wolf, I. & Gretz, N. AutoCellSeg: robust automatic colony forming unit (CFU)/cell analysis using adaptive image segmentation and easy-to-use post-editing techniques. **8** (2018).
42 Ma, H., Beiter, R., Gaultier, A., Acton, S. T. & Lin, Z. in *IEEE International Conference on Image Processing (ICIP).*

43 Esteva, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115-118, doi:10.1038/nature21056 (2017).
44 Choolani, M. *et al.* Characterization of first trimester fetal erythroblasts for non-invasive prenatal diagnosis. *Mol Hum Reprod* **9**, 227-235, doi:10.1093/molehr/gag027 (2003).
45 Bhat, N. M., Bieber, M. M. & Teng, N. N. One-step enrichment of nucleated red blood cells. A potential application in perinatal diagnosis. *J Immunol Methods* **158**, 277-280, doi:10.1016/0022-1759(93)90224-u (1993).
46 Bianchi, D. W. *et al.* Fetal gender and aneuploidy detection using fetal cells in maternal blood: analysis of NIFTY I data. National Institute of Child Health and Development Fetal Cell Isolation Study. *Prenat Diagn* **22**, 609-615, doi:10.1002/pd.347 (2002).
47 Tao, D., Shen, Y., Feng, X. & Chen, H. The application of CD71 and Hoechst33258 to staining method for sorting fetal nucleated red blood cells in the peripheral blood of pregnant women. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* **17**, 352-354 (2000).
48 Nemescu, D. *et al.* Comparison between paramagnetic and CD71 magnetic activated cell sorting of fetal nucleated red blood cells from the maternal blood. **34**, e23420, doi:10.1002/jcla.23420 (2020).
49 Bandyopadhyay, A. *et al.* Cytology Microarray on Cell Block Preparation: A Novel Diagnostic Approach in Fluid Cytology. *J Cytol* **36**, 79-83, doi:10.4103/JOC.JOC_15_17 (2019).
50 Abe, H. *et al.* Insulinoma-associated protein 1 is a novel diagnostic marker of small cell lung cancer in bronchial brushing and cell block cytology from pleural effusions: Validity and reliability with cutoff value. *Cancer Cytopathol* **127**, 598-605, doi:10.1002/cncy.22177 (2019).
51 Woo, C. G. *et al.* Diagnostic benefits of the combined use of liquid-based cytology, cell block, and carcinoembryonic antigen immunocytochemistry in malignant pleural effusion. *J Thorac Dis* **10**, 4931-4939, doi:10.21037/jtd.2018.07.139 (2018).
52 Chougrad, H., Zouaki, H. & Alheyane, O. Deep Convolutional Neural Networks for breast cancer screening. *Comput Methods Programs Biomed* **157**, 19-30, doi:10.1016/j.cmpb.2018.01.011 (2018).
53 Hussain, Z., Gimenez, F., Yi, D. & Rubin, D. Differential Data Augmentation Techniques for Medical Imaging Classification Tasks. *AMIA Annu Symp Proc* **2017**, 979-984 (2017).
54 Ganesan, P., Rajaraman, S., Long, R., Ghoraani, B. & Antani, S. Assessment of Data Augmentation Strategies Toward Performance Improvement of Abnormality Classification in Chest Radiographs. *Annu Int Conf IEEE Eng Med Biol Soc* **2019**, 841-844, doi:10.1109/EMBC.2019.8857516 (2019).

459

460 **Figure legends**

461 **Figure 1. Flow chart of fNRBCs sorting and CAD based on CNN.**

462 A . fNRBC sorting, fixation, and recognition.

463 B-a.Cellular-level ROI extraction: an accurate cell contour at different magnifications.

464 B-b. Examples of images processed via data augmentation.

465 C. Prediction network: schematic representation of the framework of the prediction
466 network.

467 **Figure 2. Gray distribution histogram (A), the precision-recall curve of p-net and**
468 **CNN networks (B), and HE staining of fNRBCs in different slices（C）.**

469 CW: clockwise.

470 **Supplementary figure 1. Structure of the CNN model.**

471 # Figures 1 and 2 need color in print.

472

Figures

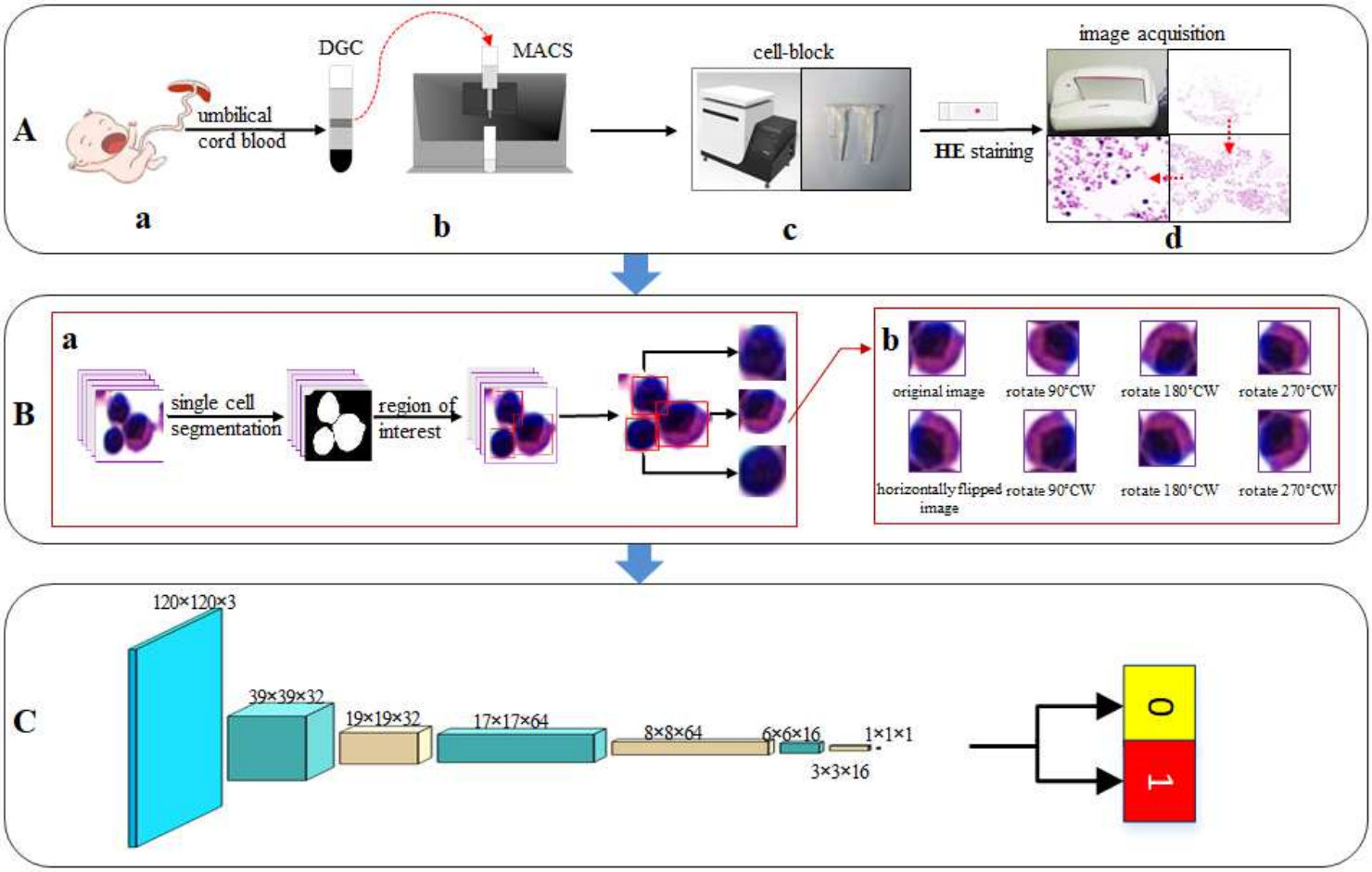

Figure 1

Flow chart of fNRBCs sorting and CAD based on CNN. A . fNRBC sorting, fixation, and recognition. B-a.Cellular-level ROI extraction: an accurate cell contour at different magnifications. B-b. Examples of images processed via data augmentation. C. Prediction network: schematic representation of the framework of the prediction network.

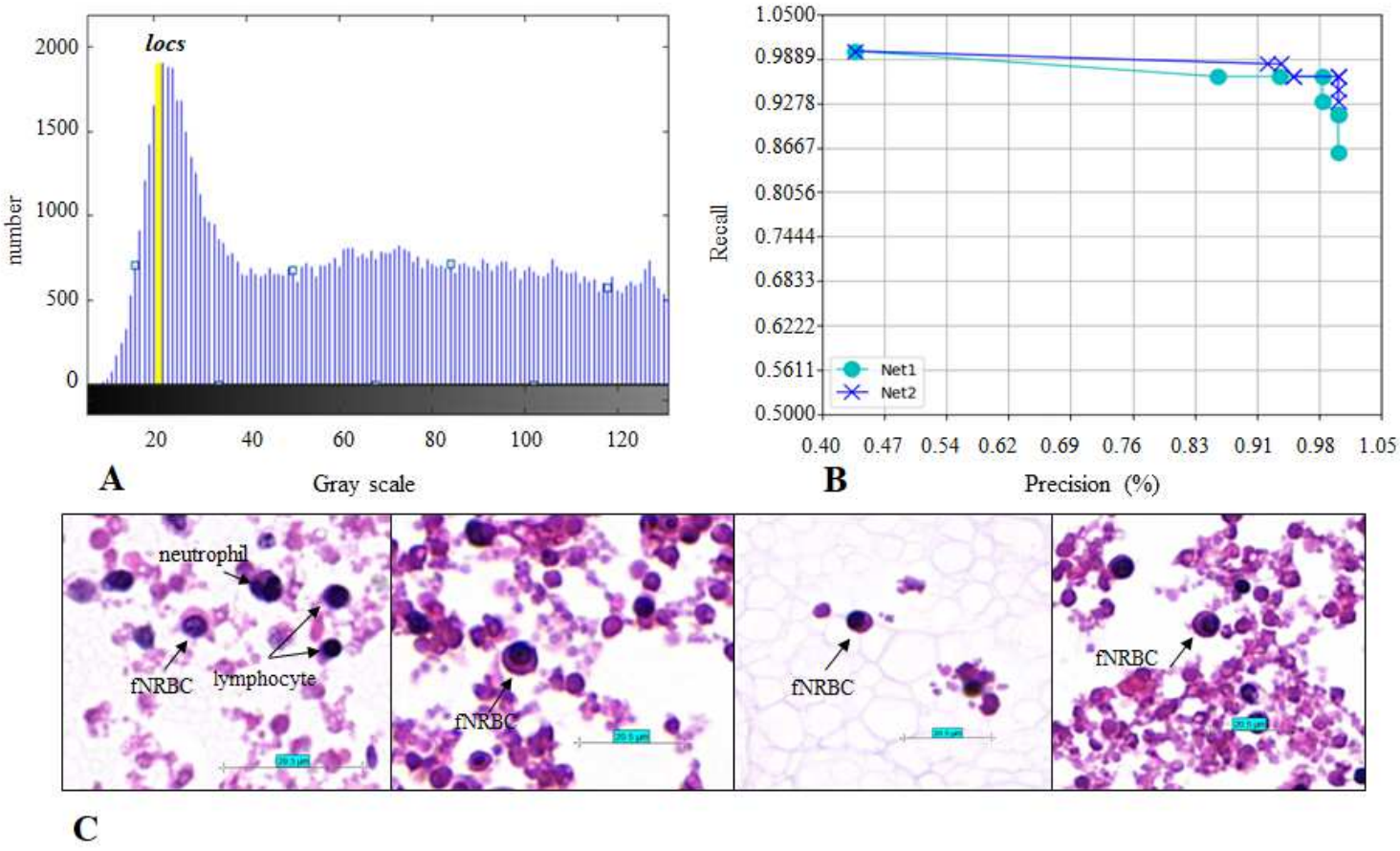

Figure 2

Gray distribution histogram (A), the precision-recall curve of p-net and CNN networks (B), and HE staining of fNRBCs in different slices（C）. CW: clockwise.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS01.png