Neurocomputing 223 (2017) 86–102

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Automatic placental maturity grading via hybrid learning

Baiying Lei[a], Ee-Leng Tan[b,1], Siping Chen[a], Wanjun Li[a], Dong Ni[a], Yuan Yao[c,*], Tianfu Wang[a,*]

[a] School of Biomedical Engineering, Shenzhen University, National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Nanhai Ave 3688, Shenzhen 518060, Guangdong, China
[b] School of Electrical and Electronic Engineering, Nanyang Technological University, 50 Nanyang Avenue, 639798, Singapore
[c] Department of Ultrasound, Affiliated Shenzhen Maternal and Child Healthcare, Hospital of Nanfang Medical University, Shenzhen 518060, China

ARTICLE INFO

Communicated by: Dr. Deng Cheng

Keywords:
Placental maturity evaluation
Pyramidal descriptor
Deep feature training
Hybrid learning
Normalization

ABSTRACT

Fetal viability, gestational age, and complicated image processing have made evaluating placental maturity a tedious and time-consuming task. Despite various developments, automatic placental maturity still remains as a challenging issue. To address this issue, we propose a new method to automatically grade placental maturity from B-mode ultrasound (BUS) and color Doppler energy (CDE) images based on a hybrid learning architecture. We also apply an improved pyramidal shift invariant feature transform (IPSIFT) descriptor using a coarse-to-fine scale representation for visual feature extraction. These local features are then clustered by a generative Gaussian mixture model (GMM) to incorporate high order statistics. Next, the clustering representatives are encoded and aggregated via Fisher vector (FV). Instead of using traditional FV, an end-to-end deep training strategy is developed to fine-tune the GMM parameters to boost evaluation performance. A multi-view fusion technique is also developed for feature complementarity exploration. Extensive experimental results demonstrate that our method delivers promising performance in placental maturity evaluation and outperforms competing methods.

1. Introduction

In recent years, the use of radiation-free and low-cost ultrasound (US) imaging to evaluate placental maturity [1,2] has gained more and more popularity. B-mode grayscale US (BUS) images have been widely used for detecting placental abnormalities (e.g., miscarriage, fetal death, still birth, small gestational age, and various pregnancy complications). Although other imaging modality is available [3], routine pregnant woman screening has largely been based on BUS and color Doppler energy (CDE) imaging [4]. Due to numerous fetal viabilities, various gestational ages, and complicated image processing, evaluating placental maturity requires knowledge and substantial prior experience of clinician to effectively perform this routine analysis. Variations of calcification degree and image quality constraints render subjective measurement unappealing as well [5]. Due to these reasons, placental maturity evaluation remains a challenging issue. Automatic computer-assisted diagnosis would be desirable as it not only reduces errors caused by judgment [6], but also provides an attractive and meaningful standardization tool [7].

The first step to develop an automatic method to evaluate placental maturity is feature representation. Determining appropriate feature detectors and descriptors is a challenging task due to intra- and inter-class variations, speckle artifacts, and low resolutions of US images [8–10]. To date, popular feature representations include single feature representation (i.e., histogram of gradient (HoG) [11], scale invariant feature transform (SIFT) [12]) and fusion of various single feature representations. Local feature descriptor such as SIFT has demonstrated its ability to find dominant and distinctive information. Apart from the traditional SIFT features, there are some enhanced or extended SIFT features such as the densely sampled scale-invariant feature transformation (DSIFT) [13–18], RootSIFT [19], the pyramidal histogram of visual words (PHOW) [15,20], pyramidal histogram of gradient (PHoG) [21], local pyramidal descriptor [22], local intensity order pattern (LIOP) [23], and DAISY [24,25]. These enhanced features achieve superior classification performance as compared to typical SIFT descriptor in the computer vision field [24,26,27]. A spatial pooling strategy is quite effective for many descriptors such as SIFT-like color descriptor [28], DAISY-like descriptor [24,25] and HoG [11]. For instance, spatial pyramid matching (SPM) [29] improves classification performance by exploring pairwise or neighboring information [9]. The essential idea of SPM is to split signal into various parts, regions and scales [30]. Inspired by the promising performance

* Corresponding authors.
E-mail addresses: yaoyuan90@126.com (Y. Yao), tfwang@szu.edu.cn (T. Wang).
[1] Present address: Beijing Sesame World Co. Ltd, Beijing, China.

http://dx.doi.org/10.1016/j.neucom.2016.10.033
Received 8 December 2015; Received in revised form 1 September 2016; Accepted 23 October 2016
Available online 27 October 2016
0925-2312/

of the visual discriminative features and SPM strategy, we believe that these strategies can effectively minimize or resolve some issues of evaluating placental maturity [31]. For instance, instead of using conventional descriptors, an enhanced feature descriptor via SPM technique was developed to boost staging performance [32]. Based on this intuition, we devise a new visual descriptor integrating SPM partitioning techniques to represent the local feature. We refer this new representation as IPSIFT, which normalizes the feature to improve grading performance.

The evaluating of placental maturity is commonly plagued by the dimensionality curse (i.e., small training and large feature dimension). This problem can be resolved by feature encoding, which explores high level information. This approach not only identifies the discriminative feature, but also reduces the feature dimension. Representative feature encoding includes linear sparse coding [33], linear locality constrained coding [34], linear locally embedding (LLE) [31], and support vector coding (SVC) [33]. Visual word occurrence representation, bag of feature (BoF) framework such as the bag of visual words (BoVW) [29,35] has attracted extensive interest in recognition task. BoVW is aggregated with soft assignment and data-dependent encoding into enhanced versions such as the vector of locally aggregated descriptor (VLAD) [16], and Gaussian mixture model (GMM) clustering [13] based Fisher vector (FV) [36]. Typically, mapping into visual dictionary and pooling into histogram representation for classification have become a hot topic in a large number of applications [19]. In addition, the semantic pooling approach is able to enhance the classification performance [37].

Shallow structure in existing BoF pipeline achieves unappealing performance, whereas deep structure motivated from deep learning has obtained more encouraging classification performance [38]. Deep feature architecture design has been widely studied for classification [39], and deep learning based representation [9] is gaining substantial popularity especially in the medical field because of its powerful feature representation ability [40]. By integrating feature weight into deep representation, such representation outperforms the standard hand-craft or binary representation [41]. Inspired by the popular deep learning method [42], we integrate a novel feature hierarchy having deep architecture to evaluate placental maturity. This deep FV strategy takes advantage of spatial information to improve the staging performance by expanding the expressive power of the feature representation.

Automatic placental maturity evaluation via US modality is one of the most commonly used methods [43,44]. It was argued that methods based on one modality are limited in terms of performance due to omission of other information such as blood flow information [3]. Blood vessels as non-branched and branched blood vessels are important in morphogenesis during pregnancy [45–47], and play an essential role in placental maturity evaluation. Meanwhile, CDE imaging provides the velocity distribution of blood flow from grayscale or color band in the human body tissue plane. CDE also can provide the structure of human tissue and the body's blood flow (or organization) kinematic information [4]. Based on these observations, we believe that CDE is an appropriate technique to boost the evaluation performance using the blood flow information. However, this information is omitted in previous methods. In view of this, we incorporate the vessel information via CDE in our placental maturity evaluation. By integrating CDE and BUS, high performance is expected since each modality complements each other. To the best of our knowledge, this is the first investigation of combining CDE and BUS for placental maturity evaluation. Our evaluation is mainly based on the benchmark tool specified in Table 1 [48,49] according to placental variations, chorionic plate, placental substance, and basal layer.

The main objective of this paper is to devise an automatic placental maturity evaluation method via a benchmark tool [1]. We propose a novel hybrid learning framework fusing CDE and BUS modalities with deep feature architecture. The main contributions of our work are four-fold: 1) CDE is integrated with BUS images for feature fusion; 2) A novel feature descriptor is designed using improved SPM strategy; 3) An end-to-end deep feature architecture is developed with multi-layer design; 4) A multi-view technique is devised to explore feature complementarity. This is the first work for a comprehensive evaluation of the placental maturity using BUS and CDE, and the empirical evidence and feasibility study of the proposed method would be presented in this paper.

2. Related work

In the literature, there are some interesting works for placental maturity staging. The first significant attempt for placental evaluation was carried by Grannum et al. in 1979, which automatically graded the placental maturity based on the chorionic, substance, and basal plates. This was the first benchmark tool for placental evaluation. Subsequently, Linares [48] et al. proposed to use textual features, laws masks, and gray-tone difference for placental maturity evaluation. This method was validated by the leave one out method, and classified by k-nearest neighboring method. However, the best grading accuracy is 60.71%, which is too low for clinical practice. To improve grading accuracy, Liu et al. [2] suggested to grade the placental US images based on calcification degree. This method utilized variance, distortions, and kurtosis as features, and support vector machine (SVM) as classifier to classify the extracted features. However, this method is not entirely automated as it requires subjective measurement. Although there are some methods to address the placental maturity evaluation issue using objective methods [2], one common shortcoming of these methods is the ignorance of the scale and rotation invariant features, and hence these methods lack comprehensive evaluation and possess unsatisfactory grading performance.

To overcome the shortcomings of prior methods and to develop desirable feature descriptors, the scale invariant feature descriptors such as SIFT, DAISY, and LIOP that deliver promising performance in a myriad of applications. For example, an US image retrieval method was proposed in [31] based on extracted features of SIFT, local binary pattern (LBP), and image intensity value. This method delivers quite promising results with publicly available dataset. Generally, effective representation is of vital significance in classification and recognition work. The popular representation starts from the low-level descriptor and builds a code book or dictionary from the training samples, which are then encoded and aggregated into a global representation. Motivated from promising performance, features are represented by a histogram of word appearance regardless of word order and their locations. The most typical BoF representations include BoVW, VLAD, sparse coding, and soft-assignment based FV. FV represents the sample with parameter gradient vector using the log likelihood. For example, the BoF representation using DSIFT was applied in medical classification task and obtained attractive performance [43]. FV had also been widely applied in placental maturity staging due to its effectiveness via high-order information exploration. For example, Lei et al. [50] suggested utilizing LIOP and FV to develop automatic placental maturity evaluation method, which demonstrated that histogram representation and feature encoding approach were highly suited for this task. Generally, FV achieves better performance in the grading task than VLAD and BOVW as FV captures high level information from the BUS image. In fact, FV incorporate not only the first order statistics, but also the second order statistics [19]. Thus, FV outperforms VLAD, which is based on only the first order statistics. Also, Li et al. [50] proposed to use densely sampled DAISY descriptor to uncover underlying placental information, and achieved satisfactory maturity staging performance. This method benefits not only from leveraging the high order information, but also from densely extracting features. In super vector encoding methods [51], high order statistics of the associate descriptor maintains the discrimination. Motivated from it, VLAD was augmented with high order information in [52], where the high order