Biomedical Signal Processing and Control 43 (2018) 18–30

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Computerized detection of spina bifida using SVM with Zernike moments of fetal skulls in ultrasound screening

Umut Konur

Computer Engineering, Bülent Ecevit University, Zonguldak, Turkey

ARTICLE INFO

Article history:
Received 8 June 2017
Received in revised form 16 January 2018
Accepted 18 February 2018

Keywords:
Spina bifida
Lemon sign
Ultrasound
Zernike moments
SVM
Rare class

ABSTRACT

A computer aided detection scheme for the neural tube defect of spina bifida is proposed. Features from Zernike moments of fetal skull regions viewed by ultrasound are utilized in SVM classification. Rotational invariance of magnitudes of Zernike moments and their easy normalization with respect to translation and scale make them attractive for image and shape description. In particular, they are perfect candidates for classifying shapes of fetal skulls that possess markers of spina bifida. The automated detection system may act in decision support to help specialists avoid false negatives. Problems of rarity are handled with combinations of oversampling and undersampling. A variant of the synthetic minority oversampling technique (SMOTE) and random undersampling (RU) have been applied on training data. Experiments show the trade-off in various performance indicators depending on different sampling choices. The average values of 0.6276 *F*-measure and 0.6306 GMRP are achieved on non-sampled (original) test sets when training is performed using sampled data after 400% borderline-SMOTE followed by 50% RU with respective accuracy and specificity realizations of 94% and 98%.

1. Introduction

*C*omputer *a*ided *d*iagnosis/*d*etection (CADx/CADe) is an interdisciplinary area aiming to play a supporting role in medicine. Digital processing of radiological images and machine learning are combined in CAD systems. Inputs comprise images of *m*agnetic *r*esonance *i*maging (MRI), *c*omputed *t*omography (CT), X-ray and *u*ltrasound (US) modalities. In CAD systems, *r*egions *o*f *i*nterest (ROI) in images are analyzed and the goal is to detect conspicuous tissues/structures. Examples of CAD include detecting malignant pulmonary nodules [1], mammographic masses [2], soft tissue tumors [3], some cancers [4,5] and many more.

In this work, a CAD system for detecting the neural pathology of *spina bifida* from fetal ultrasound (US) images has been implemented. Spina bifida (open/split spine) is a common birth defect from the category of *n*eural *t*ube *d*efects (NTD) which affects the spine and spinal cord. The stages in embryo development involve nervous system formation during which the neural plate tissue folds to form a tube which further folds into the spinal cord. The incorrect folding of the neural plate may cause spina bifida to give rise to an abnormally formed section of the spinal column. The defect may cause bladder control problems, sensation loss and paralysis. Fig. 1 shows US images of an axial spine, a sagittal spine and a transcerebellar skull, all associated with defective fetuses. The worldwide prevalence of spina bifida is one-two cases per 1000 births.

US screening and examination can let diagnose neural tube defects. Detecting spina bifida as early as possible in the prenatal stage is vital for careful planning and remedies. Besides observing the spine, markers of fetal skulls also help in the diagnosis. Checking the existence of these markers is easier for machine processing. A typical marker associated with skull shapes is the lemon sign [6] which appears when the frontal bones of a skull look flattened and inwardly bent. The malformed fetal skull in Fig. 1 is one with lemon sign.

Recognizing a defective fetus is naturally a two-class classification problem (i.e. defective (positive) or healthy (negative)). Solutions for the problem have been attempted in previous research. The methods rely on features of skull contours [7,9,10] and regions [8]. The decision rules have generally exploited the existence of lemon sign [6]. Either thresholds [7,9], SVMs [8] or nearest neighbors [10] have been used for classification.

The earliest work [7] considers segmentation as part of the problem and fits elliptical models to detect contour segments on skull boundaries. The longest of the several detected segments is assumed to be indicative of the defect. Deciding whether a fetus is defective is based on the fraction of segment points (i.e. pixels), that do not match the expected pattern of slope values of tangent lines for a healthy appearance. Absolute values (i.e. magnitudes) of Zernike moments and SVMs have been used in another work

E-mail address: konur@beun.edu.tr

https://doi.org/10.1016/j.bspc.2018.02.012
1746-8094/

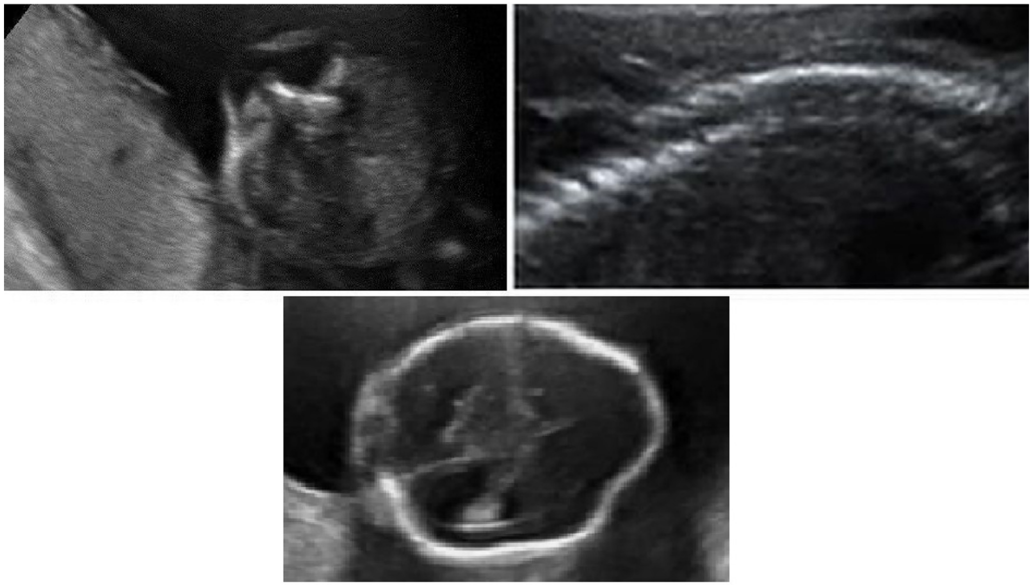

Fig. 1. Axial spine (top-left), sagittal spine (top-right) and transcerebellar skull (bottom) sonograms of fetal spines with spina bifida.

[8]. Both transcerebellar skull and axial spine images of fetuses have been considered for a more secure detector. The drawbacks include adhoc approaches for segmenting skulls [7,8] and spines [8] in addition to the uncertainty in the number of Zernike moments that should be used [8].

The latest two efforts [9,10] have benefited from the *c*urvature *s*cale *s*pace (CSS) [11–14] representation of curves to extract skull features. CSS provides descriptions of contours at multiple scales (*l*evels *o*f *d*etail (LOD)). The CSS representation of sampled contour points has allowed measuring the similarity of shapes in a framework with scale, translation and rotation invariance. Classification relies on either comparing the similarity scores of input CSS images to thresholds determined from a set of training images belonging to two distinct classes [9] or the simple nearest neighbor classifier operating on similarity scores of data samples to those in training sets [10]. The methods have been tested with the same data set utilized in this paper. This is a relatively richer data set (29 defective and 329 healthy samples). The weakness of threshold selection mechanisms and the absence of rare class handling sacrifice from robustness in the work of [9]. Combinations of oversampling and undersampling have been employed to handle class rarity and imbalance in [10]. The classification performances of nearest neighbor classifiers have been reported on a variety of settings depending on different types of CSS features (i.e. how similarity scores are computed) and utilizing only the actual contours of skulls in CSS matching or utilizing both actual contours and their reflections on a circular mirror centered at the center of gravity of actual contours [15]. Although the system has shown to perform well for some settings, one may deem that sampling is not satisfactory enough. Furthermore, the property of interpretability of a parametric classifier is missing.

The current work is based on detecting lemon sign associated with fetal skulls to decide if a fetus is defective. Lemon sign is an entity along skull contours. It is equivalently connected with shapes of skull regions. From a general viewpoint, representing 2D planar shapes can be performed with various methodologies. Zhang and Lu [16] provide a review of shape representation techniques. The shape features subject to the work of this paper are derived using Zernike moment transforms [17,18] of fetal skull shapes. Magnitudes of Zernike moments computed up to a certain order are used with an SVM (*s*upport *v*ector *m*achines) classifier. Attributable to their orthogonality property, Zernike moments have been popular for shape representation and classification. The combination of magnitudes of Zernike moments used as shape features and SVM [19–21] as classifiers is appropriate, because magnitudes of Zernike moments are rotation-invariant, translation and scale normalizations are easy and SVM is state-of-the-art.

Besides the main building blocks, the rarity of samples used to train the classifier is another issue. There may be two types of data rarity in a two-class problem. The number of available samples may be few (i.e. absolute rarity) or the members of one class may be much fewer than those of the other (i.e. relative rarity or the class imbalance problem). Such training data induce improper classifier design because available instances can not partition and represent the input space well enough. This degrades the accuracy and generalization capability of classifiers. In addition, the classical performance metric of accuracy is not satisfactory since positive (rare) and negative (frequent) class decisions are equally-favored leading to incorrect conclusions. As a matter of fact, the US image samples of this work highly reflect the properties of absolute and relative rarity. Tackling rarity is generally by modifying data distributions to represent rare classes satisfactorily in the input space. Sampling [22] is a common approach. Boosting refers to general sampling where weights on training samples are iteratively updated so that base learners can focus on examples that were misclassified in previous iterations. Sampling techniques include undersampling and oversampling. Undersampling discards some frequent (majority) class samples from the training set and oversampling adds exact copies of rare (minority) class instances to it. A more prominent oversampling approach is the *s*ynthetic *m*inority *o*versampling *te*chnique (SMOTE) [23] which generates synthetic rare class samples using each rare sample and its k nearest neighbors. Borderline-SMOTE [24] is a variant which synthesizes new samples only for those rare class samples that are prone to classification error. The CAD system exploits borderline-SMOTE and *r*andom *u*ndersampling (RU). The classification performance is reported through *r*eceiver *o*perating *c*haracteristics (ROC) analysis [25] and the associated *a*rea *u*nder the ROC *c*urve (AUC) [26] in addition to the point metrics of accuracy, recall, precision, specificity, GMRP (*g*eometric *m*ean of *r*ecall and *p*recision) and *F*-measure.

A block diagram of the proposed CAD system is presented in Fig. 2. In an ideal scheme, the working system includes an auto-