European Journal of Molecular & Clinical Medicine
ISSN 2515-8260 Volume 7, Issue 8, 2020

FALSE POSITIVE REDUCTION BASED ON ANATOMICAL CHARACTERIZATION USING DEEP LEARNING NEURAL NETWORK IN LUNG NODULE DETECTION

[1]K. S. Gowri Laksshmi and [2]Dr. R. Umagandhi
[1]Research Scholar and [2]Associate Professor and Head Department of Computer Science Department of Computer Technology, Kongunadu Arts and Science College, Coimbatore

Abstract* - *Accurate and early diagnosis of Lung Cancer increases survival rate of patient. Diagnosis of Lung Cancer involves identifying tumour as either benign or malignant, this categorizing done by the Image mining techniques called classification. Image classification is the primary domain, in which Deep neural networks play the most important role of medical image analysis. The image classification accepts the given input images and produces output classification for identifying whether the disease is present or not. Image data represents a keystone of many research areas including medicine, forensic criminology, robotics and industrial automation, meteorology and geography as well as education. In this paper proposal methodology is an integration of ensemble classification is completed using the Entropy Weighted Residual Convolution Neural Network (EWRCNN). Finally, the results are evaluated between the samples, compared to FP reduction with Faster R-CNN alone, the inclusion of rule-based classification lead to an improvement in detection accuracy for the CAD system. These preliminary results demonstrate the feasibility of the proposed EWRCNN approach to lung nodule detection and FP reduction on CT images. The objective of this research is to predict correct Lung Cancerous nodule and classifying in CT and X-ray image.

***Keyword:* lung Nodule Detection, Faster R-CNN, FP reduction, Entropy Weighted Residual Convolution Neural Network.**

INTRODUCTION

People with an increased risk of lung cancer may consider annual lung cancer screening using low-dose CT scans. Lung cancer screening is generally offered to people 55 and older who smoked heavily for many years or who have quit in the past 15 years. If there's reason to think that you may have lung cancer, your doctor can order a number of tests to look for cancerous cells and to rule out other conditions.

Tests may include:

- Imaging tests. An X-ray image of your lungs may reveal an abnormal mass or nodule. A CT scan can reveal small lesions in your lungs that might not be detected on an X-ray.
- Sputum cytology. If you have a cough and are producing sputum, looking at the sputum under the microscope can sometimes reveal the presence of lung cancer cells.
- Tissue sample (biopsy). A sample of abnormal cells may be removed in a procedure called a biopsy.

Doctor can perform a biopsy in a number of ways, including bronchoscope, in which your doctor examines abnormal areas of your lungs using a lighted tube that's passed down your throat and into your lungs. A biopsy sample may also be taken from lymph nodes or other areas where cancer has spread, such as your liver. Careful analysis of your cancer cells in a lab will reveal what type of lung cancer. Results of sophisticated testing can tell your doctor the specific characteristics of your cells that can help determine your prognosis and guide your treatment. This work carried via image mining.

Intelligently classifying image by content is an important way to mine valuable information from large image collection. The classification module in the mining system is usually called classifier. One of the main problems in computer vision is the image classification problem which is concerned with determining the presence of visual structures in an input image. Image classification analyzes the numerical properties of various image features and organizes data into categories. In recent years, many advanced classification approaches, such as artificial neural networks, fuzzy-sets, and expert systems have been widely applied for image classification, but each of them having some problems and their accuracy level is comparatively less. One of the advanced classification approaches is to use the convolution neural network (CNN) architecture which obtains successful results in solving many machine learning problems. Image classification is a complex procedure which relies on different components. Here, some of the presented strategies, issues and additional prospects of image orders are addressed. The primary spotlight will be on cutting edge classification methods which are utilized for enhancing characterization precision.

Bingbing Li, Chengdong Wu [1] proposed automatic tumour brain segmentation for treatement and diagnosis. This work carried using Deep learning neural network. The residual convolutional used to extract deep features of image. An auxiliary output is added before each upper sampling as deep supervision of the network. fused weighted cross entropy used and generalized dice loss as loss function of each output path and then a weighted sum of all losses generate the final loss of the network, and use this loss to train the network. Three auxiliary outputs are then employed as deep supervision before each upper sampling to decrease the robabilities of vanishing gradient and avoid over fitting. The proposed model is evaluated on the BraTS 2018 training dataset and achieves state-of-the-art performance, with average dice scores of 0.884, 0.806, 0.702 for the whole tumor, tumor core and enhancing tumor core,

Sergey Korole,y Amir Safiullin[2] author proposed deep 3D convolutional neural network architectures for a task of classification of brain MRI scans Demonstrated performance of the residual and plain convolutional neural networks based on the ADNI dataset which is a largest available dataset of structural MRIs of subjects with Alzheimers disease and normal controls. The proposed approach should prove useful for on the fly prediction of any given MRI scan as soon as we can automatically process incoming images for skull stripping and normalization.

M`arius Vila_ Anton Bardera_ Miquel Feixas[3] In this paper, we have presented two different approaches to quantify the information content of an image taking into account In the first approach, Entropy rate, excess entropy, and erasure entropy have been used to quantify the image information from the vicinity of pixels and, in the second approach, an information channel between image regions and histogram bins has been applied to study the difficulty of extracting the information of an image. the spatial distribution of pixels. In the first approach, entropy rate, excess entropy, and erasure entropy have been used to quantify the image information from the vicinity of pixels and, in the second approach, an information channel between image regions and histogram bins has been applied to study the difficulty of extracting the information of an image.

1. DATASET

Lung Image Database Consortium image collection (LIDC-IDRI) inclusive of diagnostic and lung cancer screening thoracic computed tomography (CT) scans with marked-up annotated lesions are taken as dataset. a web-accessible international resource is considered for development, training, and evaluation for lung cancer detection and diagnosis. It is originated by the National Cancer Institute (NCI), later developed by the Foundation for the National Institutes of Health (FNIH), This data set inclusive of 1018 cases is created by seven collaborated academic centres and eight medical imaging companies. Three categories are determined for the initial blinded-read phase by each radiologist independently reviews as "nodule > or =3 mm," "nodule <3 mm," and "non-nodule > or =3 mm". The quality of the CT image is improved by forthcoming image enhancement phase. Fig. 5 represents the data set image with 512*512 LIDC-IDRI images.

1.1EXISTING WORK

In existing work, a lung nodule detection method based on deep learning is proposed for thoracic MR images. With parameter optimizing, spatial three channel input construction and transfer learning, a Faster R-CNN network is designed to locate the lung nodule region. Then a false positive (FP) reduction scheme based on anatomical characteristics is designed to reduce FPs and preserve the true nodule. This method is tested on 142 T2-weighted MR scans from the First Affiliated Hospital of Guangzhou Medical University. The sensitivity of the existing method is 85.2% with 3.47 false positives per scan. As Faster RCNN does not consider anatomical characteristics, many FP regions exist in the detection results. To reduce FPs and preserve true nodule, an FP reduction scheme based on the anatomical characteristics of lung nodule is designed

1.2PROPOSED WORK

Proposed Entropy Weighted Residual Convolution Neural Network Method

Detecting lung nodule, the classification is completed using the Entropy Weighted Residual Convolution Neural Network (EWRCNN). As less dependent on scaling, candidate extraction could be avoided by this detection scheme. Many FP regions are existed and detected in the results as Faster R-CNN without involving anatomical characteristics. The tested result exhibits the Faster R -CNN detecting many nodules and reducing the FP regions by proposed EWRCNN. Furthermore it is focused on reducing the number of false positives, increasing algorithm sensitivity, improving and optimizing the algorithm detection by considering the different sort of nodules on variety of sizes and shapes and at last, integrating ability with the Electronic Medical Record Systems and Picture Archiving and Communication Systems.

At origin phase, the lung nodule region as the lung nodule region is not completely utilized as regressed bounding box is one of the outputs by Faster R-CNN network on the other hand, lung nodule as finer bounding box is iterated of classification and bounding box regression of prior output using the regressed bounding. It is done as the plotting the defected lung nodule and being iterated still refinement of the region as nodule detection result are attained. Fig.3 shows the model of entropy weighted residual CNN. feature extraction phase perform feature extraction as Faster R-CNN as in the first iteration, region proposals are obtained from RPN and cropping and resizing of new feature maps are done by RoI pooling layer and final class scores and bounding box regression for each class are obtained using a three-layer Fully Connected (FC) network.

The first iteration gets over on proposed region and regressed bounding box with the maximum class score are select taken as region proposal of the second iteration and being of CT images. Recalculation is not necessary for RoI pooling layer input in reusing of feature maps. In all iterations, FC network uses parameter of the three-layer. Repeat this process for forthcoming iterations also.

In iteration number as 3 taken and the feature maps extraction and each$FV = feature\ extraction(image), RP_1 = RPN(FV)$

T is set iteration number and for each iteration i, the model can be represented as

$$for\ i = 1,2,\ldots T,$$

$$\begin{cases} r_i = RoIP\ ooling(F_V, RP_i) \\ scores_i, boxes_i = FC^3(r_i) \\ H_\alpha(F_V) = S_\alpha(A_1, A_2, \ldots A_C) \\ RP_{i+1} = boxes_{i,arg\max j\in\{0\ldots C\}\, scores_{ij}} \\ loss_i = loss_{cross-entropy}(scores_i, c_i) + \\ \lambda[c_i \neq bg_i] loss_{smoothL1}(boxes_i, bg_i) \end{cases}$$

the three-layer fully connected network is denoted by FC^3, class label c_i for the proposed region RP_i for C the number of classes and bg_i denotes the ground truth bounding box of object in this region, $[c_i \neq bg_i] = 1$ if the region selected is not background inclusive of object, otherwise $[c_i \neq bg_i] = 0$, and λ is the weight parameter to balance classification loss and bounding box regression loss. Renyi's α-entropy functional estimator [13] of matrix-based and its multivariate extension are exposed in this section. For a set of classesc_i, weighted entropy for a random variable X with Probability Density Function (PDF) f(x) in a finite set X(denoted as F_V),$H_\alpha(F_V)$ is defined as:

$$wH_\alpha(F_V) = \frac{1}{1-\alpha} \log \int_x P_n F_V{}^{\alpha}(x)dx$$

The number of weights in the range values for classesc_i is represented by P_n .

Estimation of entropy, joint entropy for two or more variables directly from data without PDF estimation are carried out by Renyi's entropy in terms of the normalized Eigen spectrum of the Hermitian matrix of the projected data in Replicating Kernel Hilbert Space (RKHS). Given a collection of nCT images $\{(x_1^i, x_2^i, \ldots, x_C^i)\}_{i=1}^n$, where the superscript i the image class label, each contains $C(C \geq 2)$measurements $x_1 \in X_1, x_2 \in X_2, \ldots, x_c \in X_c$ obtained from the same realization, and the positive definite kernels $\varkappa_1 : X_1 \times X_1 \to \mathbb{R}$, a matrix-based analogue to Renyi's α-entropy among C variables can be defined as:

$$S_\alpha(A_1, A_2, \ldots A_C) = S_\alpha\left(\frac{A_1 \circ A_2 \circ \ldots \circ A_C}{tr(A_1 \circ A_2 \circ \ldots \circ A_C)}\right)$$

In above equation,$(A_1)_{ij} = \varkappa_1(x_1^i, x_1^j), (A_2)_{ij} = \varkappa_2(x_2^i, x_2^j), \ldots, (A_C)_{ij} = \varkappa_k(x_C^i, x_C^j)$ and $\circ$ denotes the Hadamard product.

Smooth L_1 norm loss is used for the bounding box regression and soft maxcross entropy loss is used for classification in training [19] and for every iteration, $loss_i$ is calculated. The total of all iterations are termed as final loss (additionally RPN loss). The final output is determined by the scores and bounding boxes of last iteration as

The final output: $\begin{cases} loss_{final} = \sum_{i=1}^{T} loss_i \\ output = scores_T, boxes_T \end{cases}$

Only one region proposal are given in above output. At origin if K region are yielded by RPN it is iterated for every proposal, and final loss is as average $loss_{final}$ of all K proposals. Execution Time and Mean Square Error (MSE) are the parameters taken.

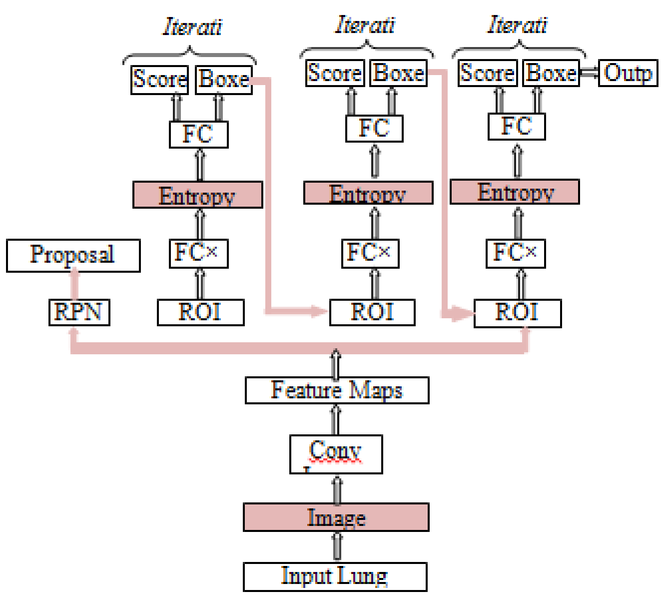


Fig.1 Architecture of Entropy Weighted Residual Convolution Neural Network

CT image: CT scans for diagnostic and lung cancer screening with marked-up annotated lesions are presented in the real time Lung Image Database Consortium image collection (LIDC-IDRI). 1018 cases in the dataset are considered in collaborated seven academic centers and eight medical imaging companies. Four experienced thoracic radiologists consider the X-ray image and its associated XML file, a clinical thoracic CT scan on each subject of for the results of a two-phase image annotation process. In starting phase of reading blindly, CT scan and marked lesions are reviewed by radiologist and said as among three categories (nodule>=3mm, nodule<3mm and non-nodule>=3mm). Fig.4 shows the result of CT-image with the nodule detection by proposed EWRCNN and existing Faster RCNN. The results evaluation results for the corresponding three methods are shown in fig 6-9. The overall results are given in Table 1 for CT.

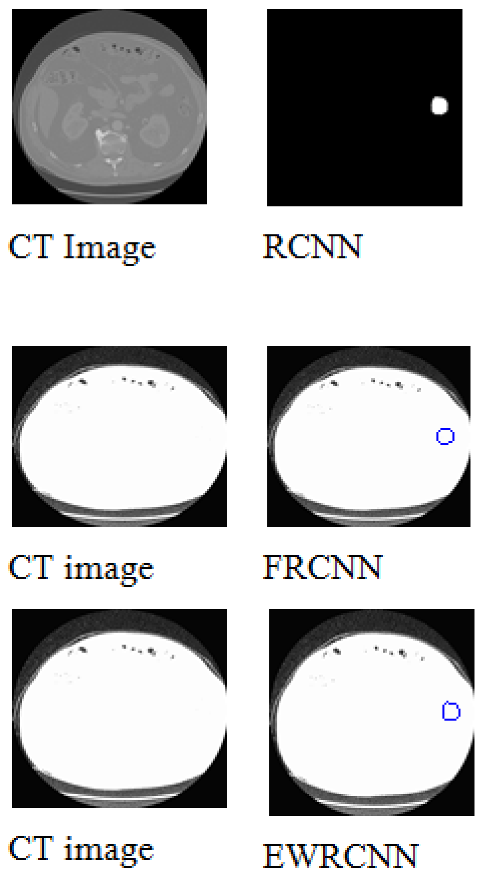


Fig.2 Lung nodule detection in CT image and its classification results

Table.1 Lung nodule detection in CT image and its classification results

	RCNN classification	Faster RCNN classification	EWRCNN Classification
Accuracy	81.0307	88.6204	91.9346
Precision	51.2075	52.2499	53.1950
Recall	81.1222	88.8299	91.3062
F- Measure	62.7836	65.7976	67.2248

Precision Result comparison **Recall Result Comparison**

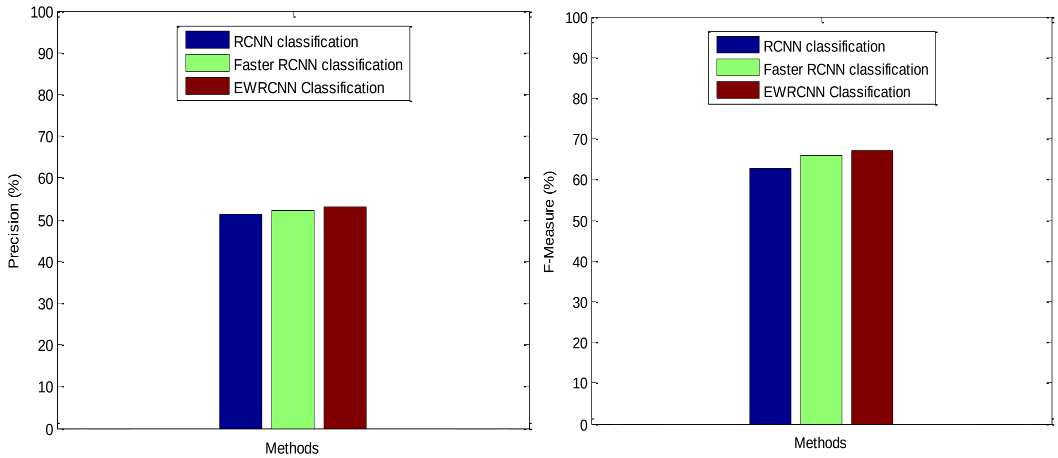


Fig.2. Comparison of compression techniques in Precision and Recall

F-measure Result Comparison **Accuracy Result Comparison**

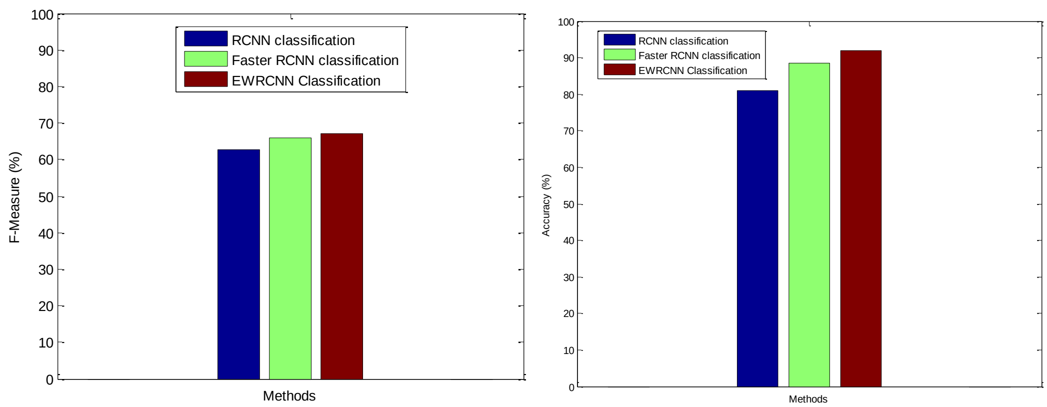
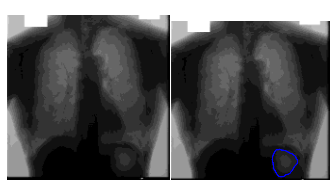


Fig.3 Comparison of compression techniques in Accuracy

The accuracy comparison results between proposed EWRCNN, and existing Faster RCNN and RCNN are shown in Fig.7. From the results, the proposed method attains high accuracy against existing methods. Efficient way of getting the lung nodule accuracy is reached as 91.9346 in proposed. When comparing the accuracy among the existing methods such as Faster RCNN and RCNN providing less rate of 88.6204 and 81.0307 respectively. In above experimental results, the proposed work gives better against existing methods.

X-ray image: exhibits the results of real time X-ray image with the nodule detection. The results evaluation results for the corresponding three methods such as proposed EWRCNN, existing Faster RCNN and RCNN are shown in fig 9-11. Figure predicts that proposed system achieves higher efficiency of precision, f-measure and accuracy. The total results of X-ray images are tabulated in Table 2.



X-ray Image Segmentation result

Table 2. Lung nodule detection of X-ray image with overall classification results for

	RCNN classification	Faster RCNN classification	EWRCNN Classification
Accuracy	82.1697	88.8420	91.7606
Precision	54.0634	56.8071	59.1270
Recall	82.2910	88.7729	91.8621
F- Measure	65.2554	69.2805	71.9460

CONCLUSION

An initial step of contrast enhancement is mandatory because of image processing techniques leading in undesired results as original image are in low contrast. Enhanced image is seen with the lung area has crystal clear view on its background. With this motivation of work based on deep learning, a lung nodule detection method is developed for thoracic MR images. Here, the image enhancement is performed initially using Fuzzy Rule based Contrast Limited Adaptive Histogram Equalization (FRCLAHE) and then feature extraction is conducted using the Fuzzy Continuous Wavelet Transform (FCWT) and gray level feature extraction (GLCM). After this step for detecting lung nodule, the classification is completed using the Entropy Weighted Residual Convolution Neural Network (EWRCNN). As less dependent on scaling, candidate extraction could be avoided by this detection scheme. Many FP regions are existed and detected in the results as Faster R-CNN without involving anatomical characteristics. The tested result exhibits the Faster R -CNN detecting many nodules and reducing the FP regions by proposed EWRCNN. Furthermore it is focused on reducing the number of false positives, increasing algorithm sensitivity, improving and optimizing the algorithm detection by considering the different sort of nodules on variety of sizes and shapes and at last, integrating ability with the Electronic Medical Record Systems and Picture Archiving and Communication Systems. Furthermore this analysis is enhanced to develop methodology to overwrite the above said pitfalls

REFERENCES

[1] Bingbing Li1, Chengdong Wu2, Jianning Chi2, Xiaosheng Yu2, Gang Wang3. A DEEPLY SUPERVISED CONVOLUTIONAL NEURAL NETWORK FOR BRAIN TUMOR SEGMENTATION. Proceedings of the 39th Chinese Control Conference July 27-29, 2020, Shenyang, China.

[2] A. Beck and M. Teboulle, “A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems,” SIAM J. Imaging Sci., vol. 2, no. 1, pp. 183–202, 2009

[3] C. Li, G. Zhu, X. u, et al., “False-Positive Reduction on Lung Nodules Detection in Chest Radiographs by Ensemble of Convolution Neural Networks,” IEEE Access, vol. 6, pp. 6 6 -7, 2018.

[4] Dimitris Perdios, Adrien Besson_, Philippe Rossinelli_, and Jean-Philippe Thiran. LEARNING THE WEIGHT MATRIX FOR SPARSITY AVERAGING IN COMPRESSIVE IMAGING This work was supported in part by the Ultrasound Togo RTD project (no. 20NA21 145911), funded by Nano-Tera.ch. IEEE 2017.

[5] M. Fadili and J.-L. Starck, “Monotone operator splitting for optimization problems in sparse recovery,” in 16th IEEE Int.Conf. Image Process., 2009, pp. 1461–1464.

[6] Jin, X., Zhang, Y., & Jin, Q. (2 6) “Pulmonary odule Detection Based on CT Images Using Convolution neural network.” 2 6 the International Symposium on Computational Intelligence and Design (ISCID). DOI: 10.1109/ISCID.2016.1053.

[7] A. Mousavi, A. B. Patel, and R. G. Baraniuk, “A deep learning approach to structured signal recovery,” in 2015 53rd Annu. Allert. Conf. Commun. Control. Compute, 2015, pp.1336–1343.

[8] M`arius Vila_ Anton Bardera_ Miquel Feixas_ Philippe Bekaert† Mateu Sbert_ ANALYSIS OF IMAGE INFORMATIVENESS MEASURES, This work has been funded in part by grants from

the Spanish Government(TIN2013-47276-C6-1-R), from the Catalan Government (2014-SGR-1232 and CONES-2010-0027), and from EU Project 610005 "Dreamspace: A Platform and Tools for Collaborative Virtual Production", call FP7 ICT- 2013-10.

[9] Roy, T., Sirohi, &Patle, A. "Classification of lung image and nodule detection using fuzzy inference system." International Conference On Computing, Communication & Automation. DOI:10.1109/CCAA.2015.7148560.

[10] Sergey Korolevy Amir Safiullin?y Mikhail Belyaevy Yulia Dodonova RESIDUAL AND PLAIN CONVOLUTIONAL NEURAL NETWORKS FOR 3D BRAIN MRI CLASSIFICATION 978-1-5090-1172-8/17/$31.00 ©2017 IEEE

[11] Yuan Chao, Min Dai, Kai Chen, Ping Chen, Zhisheng Zhang. FUZZY ENTROPY BASED MULTILEVEL IMAGE THRESHOLDING USING MODIFIED GRAVITATIONAL SEARCH ALGORITHM This work is supported in part by the National Natural Science Foundation of China (Grant No. 50805023, Grant No. 51275090), the Special Fund of Jiangsu Province for the Transformation of Scientific and Technological Achievements (Grant No. BA2010093 .978-1-4673-8075-1/16/$31.00@2016 IEEE.