2014 IEEE 27th International Symposium on Computer-Based Medical Systems

Evaluation of Classifiers to a Childhood Pneumonia Computer-aided Diagnosis System

Rafael T. Sousa
Instituto de Ciências Exatas e da Terra
Universidade Federal de Mato Grosso (UFMT)
Barra do Garças, Brazil
Email: rafaelts777@gmail.br

Oge Marques
Department of Computer & Electrical Engineering and Computer Science
Florida Atlantic University (FAU)
Boca Raton, USA

Gabriela T. F. Curado, Ronaldo M. da Costa,
Anderson S. Soares, Fabrízzio A. A. M. N. Soares,
Leandro L. G. de Oliveira
Instituto de Informática
Universidade Federal de Goiás (UFG)
Goiânia, Brazil

***Abstract*—This work extends *PneumoCAD*, a Computer-Aided Diagnosis system for detecting pneumonia in infants using radiographic images [1], with the aim of improving the system's accuracy and robustness. We implement and compare five contemporary machine learning classifiers, namely: Naïve Bayes, K-Nearest Neighbor (KNN), Support Vector Machines (SVM), Multi-Layer Perceptron (MLP) and Decision Tree, combined with three dimensionality reduction algorithms: Sequential Forward Selection (SFS), Principal Component Analysis (PCA) and Kernel Principal Component Analysis (KPCA). Current results demonstrate that Naïve Bayes classifier combined with KPCA produces the best overall results.**

I. INTRODUCTION

Pneumonia is an epidemic disease characterized by acute lower respiratory infection, usually caused by viruses or bacteria and, less commonly, other microorganisms. According to the World Health Organization (WHO), pneumonia is the leading cause of death in children worldwide, killing an estimated 1.2 million children under five years old every year. This number is higher than the mortality rate for several other diseases, such as AIDS, malaria and tuberculosis, combined [2].

Currently the best and most widely accepted imaging modality for detecting pneumonia is chest radiographs [3]. However, according Young [4] studies, errors are common in the interpretation of chest radiographs, due to inter-observer variation. This limitation of human expert-based diagnosis has provided a strong motivation for the use of computer technology to improve the speed and accuracy of the detection process.

A Computer-Aided Diagnosis (CAD) software can be defined as a second opinion in a diagnostic [5]. This kind of software is utilized to improve diagnostic accuracy, not as a means of replacing the specialist, but instead working like a second one, which is invariant to many factors that can affect the radiologist's diagnosis, such as eyestrain, distraction, stress and others.

In this work we use the features and dataset employed in previous studies [1] [6] [7], which have resulted in a full CAD system for pneumonia detection called *PneumoCAD*, which has been applied to assist in diagnostics, as well as to train and improve radiologists' expertise in childhood pneumonia detection using chest radiographs. *PneumoCAD* is currently in prototype stage. The ultimate goal behind of our system is to create a website that will provide remote diagnosis functionality by analyzing uploaded chest radiographs and processing them using image processing and machine learning algorithms.

This work was geared towards a comparative performance analysis of state-of-art classifiers combined with features selection algorithms, to improve *PneumoCAD* accuracy and find out the best classifier for childhood pneumonia detection.

A. Selected Classifiers

In this paper we apply five different classifiers, namely: The k-nearest neighbor classifier (kNN), which was used originally in our CAD system, Naïve Bayes probabilistic classifier, non-linear Support Vector Machine (SVM), neural network of Mult-layer Perceptron, and the decision tree C4.5.

II. METHODS

The images dataset used in our CAD system consists of 156 8-bit grayscale images obtained with a digital camera, that captured the chest X-rays images at a resolution of 1024 $\times$ 768 pixels. Out of these images, 78 show pneumonia while the remaining 78 do not. These images were analyzed by two trained radiologists according to WHO guidelines [8] which produced the ground truth needed to test the machine learning classifiers used in this work. The radiologists diagnosis was only considered as valid when they agree among themselves.

All features tested are based on texture and extracted in nine subspaces of Haar wavelet, like our previous paper [7].

1063-7125/14 $31.00 © 2014 IEEE
DOI 10.1109/CBMS.2014.98

TABLE I. CLASSIFIERS ACCURACY (%)

	NB	RBF-SVM	Poly-SVM	KNN	J48	MLP
SFE	65	80	69	72	62	67
PCA	56	67	66	66	56	55
KPCA	96	89	90	93	87	87

TABLE II. CLASSIFIERS AUC

	NB	RBF-SVM	Poly-SVM	KNN	J48	MLP
SFE	0.648	0.798	0.668	0.726	0.628	0.720
PCA	0.572	0.674	0.880	0.671	0.897	0.896
KPCA	0.959	0.937	0.880	0.958	0.897	0.896

All tests was made with Matlab along with their basic Toolboxes, Matlab Toolbox for Dimensionality Reduction, and Weka [9].

We performed a 10-fold cross-validation test with each classifier. Those who have parameters to be adjusted, were calibrated with a exhaustive search, testing many possible values for each parameter.

A. Dimensionality Reduction

Based on previous tests with whole feature vector, which result in a insufficiently method (70% correct rate with KNN), we decided to improve our results performing a dimensionality reduction, removing redundant and insignificant features for classification.

All classifiers was tested with each of dimensionality reduction algorithm, which are: Sequential Forward Selection (SFS). Principal component analysis (PCA), and the Kernel PCA (KPCA). KPCA tests was made with Gaussian kernel and both PCA and KPCA with 13 new dimensions.

B. Classifiers Evaluation

All tests made was evaluated with Accuracy (correct rate) to compare the overall results and AUC (Area Under Curve ROC calculated by trapezoidal approximation), which has been shown as a better measure to evaluate machine learning classifiers [10].

III. RESULTS

Accuracy results of each classifier with all three dimensionality reduction algorithms, SFE, PCA and KPCA is shown in Table I.

Table II show the graphical comparison of AUC results on each test.

The graphs and tables expose clearly the superior performance of KPCA applied with any classifier tested, with high accuracy and AUC rates, specifically with KNN and NB, where produces some good ROC curves, with area higher than 0.95. PCA do not have good results, proving to be insufficient to solve our problem. SFE have some reasonable results both in accuracy and AUC.

So the best combination found for the problem is a Naïve Bayes or K-Nearest Neighbor classifier using a feature of 13 dimensions produced with Gaussian Kernel PCA, from 17 Haralick texture features in 9 subspaces of Haar Wavelet, which provide a AUC higher than 0.95 and a accuracy of 96% (NB) and 93% (KNN). What is higher than Radiologists [4], how we can see in Table III.

TABLE III. DIAGNOSTIC ACCURACY.

Medical resident	Radiologist	NB with KPCA
66	87	96

IV. CONCLUSION

In this paper, five contemporary machine learning classifiers (Support Vector Machine, K-Nearest Neighbors, Naïve Bayes, Multi-Layer Perceptron and Decision Tree) were tested to identify and classify radiographic images in order to to detect and diagnose childhood pneumonia. The classifiers have been evaluated with a dataset taken from clinical routine. The classifiers were optimized, and tested with a cross-validation method to ensure that there is no overfitting. Naïve Bayes and K-Nearest Neighbor have shown best results (96% and 93%, respectively).

In summary, the Naïve Bayes classifier produced most accurate results and has shown to be more stable with this type of images so far. Moreover, it outperforms the best result from previous work, and even outperforms the diagnosis accuracy of Radiologists. Our future objective is test the CAD system in a real environment and evaluate if will increase medical diagnostic accuracy.

REFERENCES

[1] L. L. G. Oliveira, S. A. e Silva, L. H. V. Ribeiro, R. M. de Oliveira, C. J. Coelho, and A. L. S. S. Andrade, "Computer-aided diagnosis in chest radiography for detection of childhood pneumonia." *I. J. Medical Informatics*, vol. 77, no. 8, pp. 555–564, 2008.

[2] WHO, "Pneumonia, fact sheet n331," World Health Organization, Tech. Rep., 2012.

[3] ——, "Standardization of interpretation of chest radiographs for the diagnosis of pneumonia in children," World Health Organization: Department of Vaccines and Biologicals, Tech. Rep., 2001.

[4] M. Young and T. J. Marrie, "Interobserver variability in the interpretation of chest roentgenograms of patients with possible pneumonia," *Arch Intern Med*, vol. 154, pp. 2729–32, 1994.

[5] K. Doi, H. MacMahon, S. Katsuragawa, R. M. Nishikawa, and Y. Jiang, "Computer-aided diagnosis in radiology: potential and pitfalls," *European Journal of Radiology*, vol. 31, no. 2, pp. 97 – 109, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0720048X99000169

[6] S. O. d. Macedo and L. L. G. d. Oliveira, "Desenvolvimento de um sistema de auxílio ao diagnóstico de pneumonia na infância utilizando visão computacional," in *Workshop de Visão Computacional*, 2012.

[7] R. T. Sousa, O. Marques, F. A. A. M. N. Soares, I. I. G. Sene., L. L. G. de Oliveira, and E. S. Spoto, "Comparative performance analysis of machine learning classifiers in detection of childhood pneumonia using chest radiographs." in *ICCS*, 2013, pp. 2579–2582.

[8] T. Cherian, E. Mulholland, J. Carlin, H. Ostensen, R. Amin, M. de Campo, D. Greenberg, R. Lagos, M. Lucero, S. Madhi, K. OBrien, S. Obaro, and M. Steinhoff, "Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies," *Bull. World Health Organ.*, vol. 83(5), pp. 353–359, 2005.

[9] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *SIGKDD Explor. Newsl.*, vol. 11, no. 1, pp. 10–18, Nov. 2009. [Online]. Available: http://doi.acm.org/10.1145/1656274.1656278

[10] J. Huang and C. Ling, "Using auc and accuracy in evaluating learning algorithms," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 17, no. 3, pp. 299–310, 2005.