Alexandria Journal of Managerial Research & Information Systems

مجلة اسكندرية للبحوث الادارية ونظم المعلومات

Print ISSN : 2974-4318 online ISSN:2974-4326

PNEUMONIA DISEASE DETECTION BY BI-OBJECTIVE SUPPORT VECTOR MACHINE (BO-SVM) ON DEEP FEATURES

Hager Ali Yahia
Communication and Electronics Engineering
Alexandria, Egypt
hager2007_yahia@yahoo.com

***Abstract*—A support vector machine (SVM) learns the decision surface from two different classes of the input points, in many applications there are misclassifications in some of the input points. In this paper a bi-objective quadratic programming model is utilized and different feature quality measures are optimized simultaneously using the weighting method for solving our bi-objective quadratic programming problem. The experimental results, give evidence of the effectiveness of the weighting parameters on reducing the misclassification between two classes of the input points. The main contributions of this paper include constructing a system of a bi-objective support vector machine (BO-SVM) plus deep convolutional neural networks (CNNs)for detection the pneumonia disease using X-ray images.**

Keywords—Support vector machine (SVM); Weighting method; Quadratic programming; Deep features; Pneumonia.

I. Introduction

Support Vector Machines (SVMs) are a classification technique developed by Vapnik at the end of ’60s [1]. The theory of support vector machines (SVMs) is a new classification technique and has drawn much attention on this topic in recent years [6]. In this paper, the proposed bi-objective support vector machine is used to detect the pneumonia disease using the deep features of convolutional neural networks (CNN_s).

SVMs are known as maximum margin classifiers and find the optimal hyperplane between two classes, defined by a number of support vectors [4].

In this paper, the idea is to get all efficient solutions for the classification problem by applying the multi-objective programming technique with minimum errors and using the weighting method to solve the proposed multi-objective programming model. The remainder of this paper is organized as follows. Section 2 describes a brief review for the SVM. Section 3 describes the proposed multi-objective model for the Support Vector Machine. NEXT, section 4 describes the dataset and the system of the bi-objective support vector machine (BO-SVM) plus convolutional neural networks (CNN_s) for detection the pneumonia disease. Section 5 describes the conclusion.

II. SUPPORT VECTOR MACHINES

SVM is an efficient classifier to classify two different sets of observations into their relevant class as shown in figure 1 where there are more than straight line separates between the two sets.

The best hyperplane is the one that maximizes the margin [2].

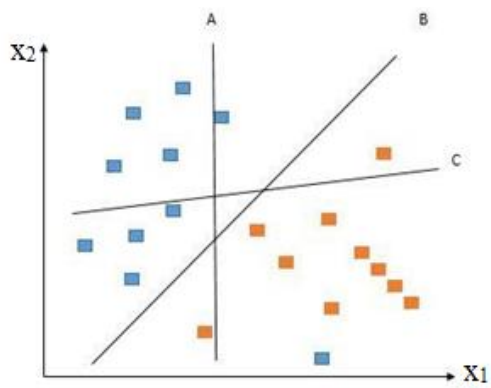
SVM has penalty parameters, and kernel parameters that have a great influence on the performance of SVM [3]. We review the basis of the theory of SVM in classification problems [7].

Let a set S of labelled training points

$$(y_1, x_1)\ldots$$
$$(y_l, x_l). \qquad (1)$$

Where, $x_i \in \mathcal{R}^N$ belongs to either of two classes and is given a label $y_i = \{-1,1\}$ for $i = 1, \ldots, l$.

Fig.1. Data classification using support vector machine



In some cases, to get the suitable hyperplane in an input space, mapping the input space into a higher dimension feature space and searching the optimal hyperplane in this feature space.
Let z = $\varphi(x)$ denote the corresponding feature space vector with mapping φ from $\mathcal{R}^N$ to a feature space z. We wish to find the hyperplane

$$w.z + b = 0 \qquad (2)$$

defined by the pair (w, b) according to the function

$$f(x_i) = sign(w.z_i + b) = \begin{cases} 1, & if\, y_i = 1 \\ -1, & if\, y_i = -1 \end{cases} \qquad (3)$$

where $w \in z$ and $b \in \mathcal{R}$. For more precisely the equation will be

$$\begin{cases} (w.z_i + b) \geq 1, & if\, y_i = 1 \\ (w.z_i + b) \leq -1, & if\, y_i = -1, \end{cases} i = 1, \ldots, l \qquad (4)$$

For the linearly separable set S, a unique optimal hyperplane can be found for which the margin between the projections of the training points of two different classes is maximized.

The optimal hyperplane problem is then regarded as the solution to the problem

$$minimize \frac{1}{2} w.w + C \sum_{i=1}^{l} \xi_i$$
$$subject\ to\ y_i(w.z_i + b) \geq 1 - \xi_i\ ,$$
$$(5)$$
$$i = 1, \ldots, l$$
$$\xi_i \geq 0,$$
$$i = 1, \ldots, l$$

where, C is a constant. The parameter C can be regarded as a regularization parameter [5]. SVM algorithms use a set of mathematical functions that are defined as the kernel.

III. THE BI-OBJECTIVEQUADRATIC PROGAMMING MODEL OF SVM

In this section, the formulation of the bi-objective programming model for the SVM plus deep convolutional neural network is described. Due to the nonlinearity separable in some of the input data, there is an error in measuring the amount of misclassification. This leads to add another objective function for the previous model (equation 5) to be in the form

$$\text{Min } \| w \|^2,$$
$$\text{Min } \sum_{i=1}^{l} \xi_i$$
$$\text{Subject to} \qquad (6)$$
$$y_i(w.x_i + b) \geq 1 + \xi_i \quad , i = 1,2, \ldots, l$$
$$\xi_i \geq 0 \quad ,$$
$$i = 1,2, \ldots, l$$

This problem is a bi-objective quadratic programming problem. The first objective is to maximize the gap between the two hyperplanes which is used to classify the input points. The second objective is to minimize the errors in measuring the amount of misclassification in case of nonlinearity separable input points.
Problem 6 can solved by the weighting method to get the set of all efficient solutions for the classification problem [9].

A. The weighting method

In this method each objective $f_i(X), i = 1,2, \ldots, k$, is multiplied by a scalar weigh $w_i \geq 0\ and \sum_{i=1}^{k} w_i = 1$. Then, the k weighted objectives are summed to form a weighted-sums objective function [8].

$$Assume\ W \text{ as} \begin{Bmatrix} w \in R^k : w_i \geq 0, \\ i = 1,2,\dots,k \\ and \sum_{i=1}^{k} w_i = 1 \end{Bmatrix} \quad (7)$$

be the set of nonnegative weights. Then the weighting problem is defined as:

$$P(W): Min \sum_{i=1}^{k} w_i f_i$$

Subject to $M = \begin{Bmatrix} X \in R^n : g_r(X) \leq 0, \\ r = 1,2,\dots,m \end{Bmatrix}.$ (8)

Then, in this paper the weighting method takes the form [9]

$$\text{Inf } z = w_1 \parallel w \parallel^2 + w_2 \sum_{i=1}^{l} \xi_i$$

Subject to

$$y_i(w.x_i + b) \geq 1 + \xi_i \quad , i = 1,2,\dots,l$$

$$\xi_i \geq 0 \quad , i = 1,2,\dots,l \quad (9)$$

$$w_1 > 0, w_2 \geq 0$$

$$w_1 + w_2 = 1$$

IV. DATASET DESCRIPTION

The Chest X-ray dataset used is publicly accessible on the Kaggle website [10], consisting of 5,863 frontal chest X-ray images (JPEG) and 2 (Pneumonia/Normal) categories. All radio-graph images in the dataset have a resolution of 1024 by 1024. Of these images, 1341 images have been identified as having pneumonia. To complement the binary classification dataset, 1341 regular X-ray images (labelled 'No Findings') were selected from the dataset.

Before being granted input to the network, the images were downscaled from 1024 by 1024 resolution to 224 by 224 resolution. Figure 2 and Figure 3 represent a part of the dataset.

Fig.2. Images for the Normal cases

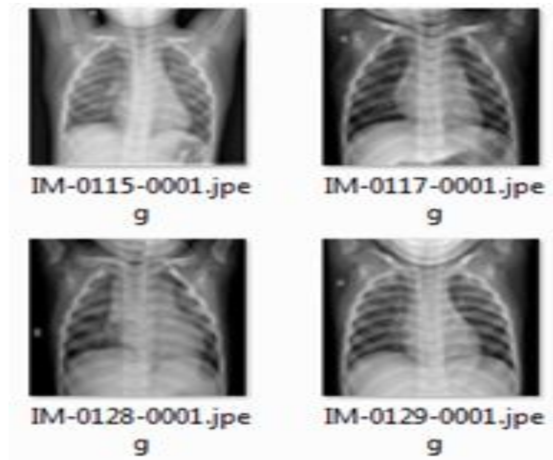
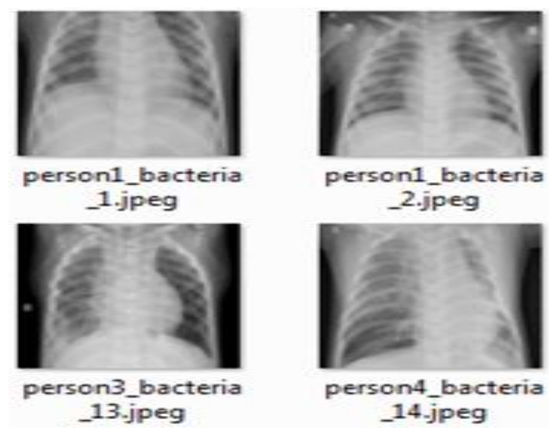


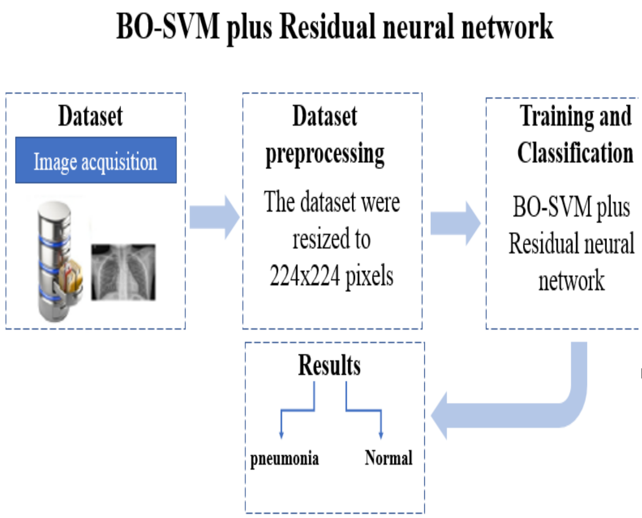
Fig.3. Images for the Pneumonia cases



A. METHODOLOGY OF THE PROPOSED MODEL

The proposed pneumonia detection system using the ‘Residual Neural Network’ (Resnet-50) is described in Figure 4. The architecture of the proposed model has been divided into three different stages - the pre-processing stage, the feature-extraction stage and the classification stage.

Fig.4. A flow diagram of the proposed system



1. Step of Pre-Processing

In most image classification tasks, the primary objective of using the Convolutionary Neural Network is to reduce the model's computational complexity, which is probably increased if the inputs are images. To minimize heavy computation and for quicker processing, the initial 3-channel images were resized from 1024 to 1024 to 224 to 224 pixels. All of the other techniques were applied to these downsized pictures [11].

2. Step of Feature-Extraction

Although the features were extracted with various variants of pre-trained CNN models, the statistical results were obtained by ResNet-50 as the optimal model for the extraction process of the feature. This process is therefore concerned with the definition of the model architecture of ResNet-50 and its contribution to the extraction of features.

3. The Classification Stage

Following feature extraction, the classification task was performed using the Bi-Objective Support Vector Machine (BO-SVM) [9]. So, with the BO-SVM classifier, the best proposed model features extracted from ResNet-50 were used to achieve better performance.

V. EXPERIMENTAL RESULTS

By using Matlab2019b program, the confusion matrix describes how the different values of the weighting parameters affect performance of ResNet-50 Model with BO-SVM and improve the accuracy.

Fig.5. $w_2 = \frac{1}{2}, w_1 = \frac{1}{2}$

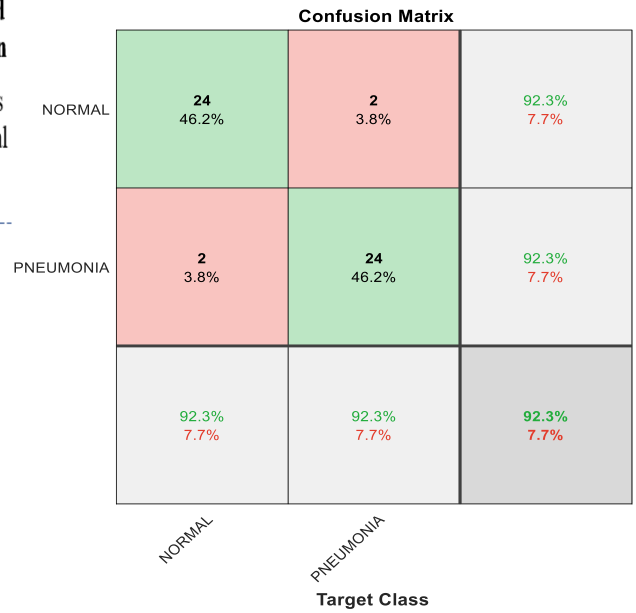


Fig.6. $w_2 = \frac{20}{21}, w_1 = \frac{1}{21}$

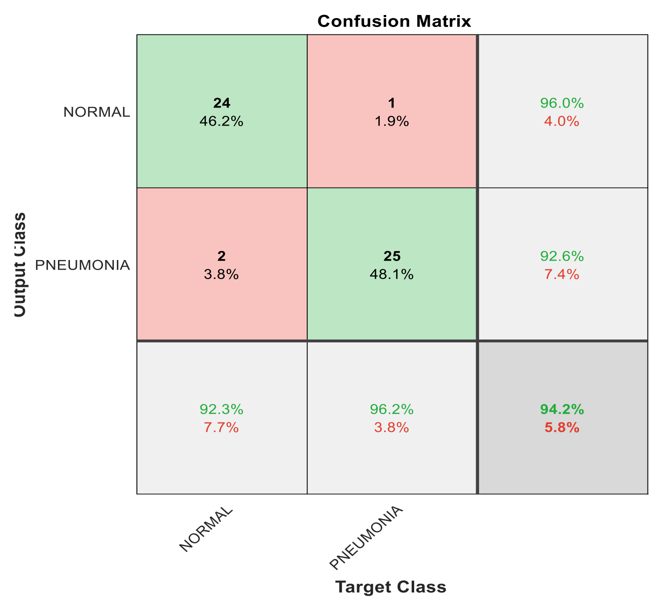
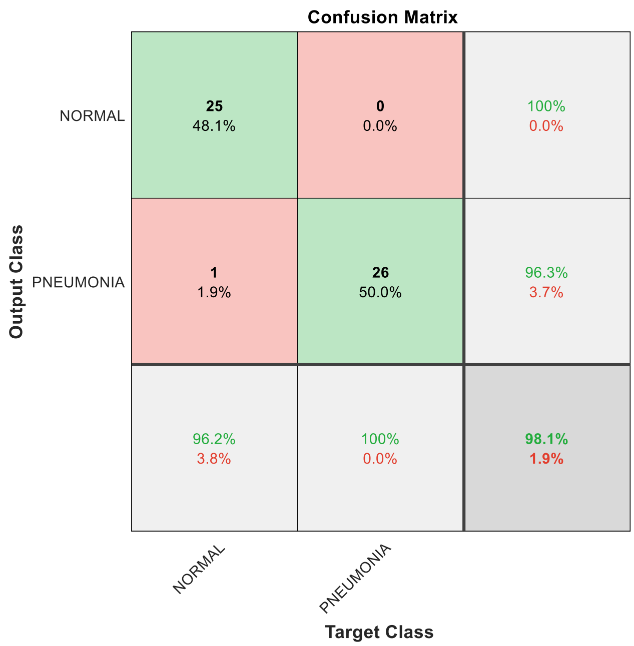


Fig.7. $w_2 = \frac{99}{100}, w_1 = \frac{1}{100}$



So, the previous results, by using different values of weighting parameters of BO-SVM, show how these parameters effect on the performance of ResNet-50 Model. For the first values of $\boldsymbol{w_1}$& $\boldsymbol{w_2}$ there are 2 Normal wrongly predicted as pneumonia and 2 pneumonia wrongly predicted as Normal as shown in figure 5. For the second values, there are 2 Normal wrongly predicted as pneumonia and 1 pneumonia wrongly predicted as Normal as shown in figure 6. Finally, for the third values, there are 1 Normal wrongly predicted as pneumonia and 0 pneumonia wrongly predicted as Normal as shown in figure 7. So, when the weighting parameter $\boldsymbol{w_2}$ is increased the performance of the Model is improved.

VI. CONCLUSION AND FUTURE WORK

This paper introduced the multi-objective programming technique for developing the set of all efficient solutions for the classification problem with minimum errors and how to solve the proposed multi-objective programming model by using the weighting method. The experimental evaluation was carried out using the dataset of Chest X-Ray publicly available on the Kaggle. The experimental results show the effect of the weighting parameters on the performance of the Residual neural network model with the bi-objective support vector machine. Finally, by changing the values of the weighting parameters, the accuracy becomes 92.3%, 94,2% then 98.1%, so the accuracy is increased by increasing $\boldsymbol{w_2}$ that associated with ξ_i (the degree of misclassification).

Our future work, building a convolutional neural network model with a fuzzy bi-objective support vector machine.

ACKNOWLEDGEMENT

At first, praise is to Allah. I would like to express my deepest gratitude and appreciation to my family who encourage me to finish this work.

REFERENCES

[1] Cortes, Corinna; Vapnik, Vladimir N (1995) "Support vector networks" (PDF). Machine
learning. 20 (3):273297. CiteSeerX 10.1.1.15. 9362. DOI:10.1007/BF00994018.

[2] Asmaa Hamad1,3(B), Essam H. Houssein1,3, Aboul Ella Hassanien2,3, and Aly A. Fahmy2 :Hybrid Grasshopper Optimization Algorithm and Support Vector Machines for Automatic Seizure Detection in EEG Signals. Faculty of Computers and Information, Minia University, Minya, Egypt. January 2018. DOI: 10.1007/978-3-319-74690-6_9.

[3] Alaa Tharwat1;_, Thomas Gabel1, Aboul Ella Hassanien2;_ Parameter Optimization of Support Vector Machine using Dragon_y Algorithm. Faculty of Computer Science and Engineering, Frankfurt University of Applied Sciences, Frankfurt am Main, Germany ,Faculty of Computers and Information, Cairo University, Egypt. January 2018 DOI: 10.1007/978-3-319-64861-3_29.

[4] Gray, D., Bowes, D., Davey, N., Sun, Y., Christianson, B.: Using the Support Vector Machine as a Classification Method for Software Defect Prediction with Static Code Metrics. In: Palmer Brown, D., Draganova, C., Pimenidis, E., Mouratidis, H. (eds.) EANN 2009. Communications in Computer and Information Science, vol. 43, pp. 223–234. Springer,Heidelberg (2009).

[5] Chun-Fu Lin and Sheng-De Wang: Fuzzy Support Vector Machines. Article in IEEE Transaction on neural networks March 2002. DOI:10.1109/72.991432.

[6] C. Burges, A tutorial on support vector

machines for pattern recognition, Data Mining and Knowledge Discovery, vol.2, no.2,1998.
[7] C. Cortes and V. N. Vapnik, “Support vector networks,” Machine Learning, vol.20, pp.273-297,1995.
[8] Chankong V. and Haimes Y.Y., Multi-objective Decision-Making: Theory and Methodology (North Holland Series in System Science and Engineering, 1983).
[9] Mohammed Zakaria Moustafa, Mohammed Rizk Mohammed, Hatem Awad Khater and Hager Ali Yahia, Building a Bi-objective Quadratic Programming Model for The Support Vector Machine, 8th International Conference on Artificial Intelligence, Soft Computing (AISC 2020), DOI: 10.5121/csit.2020.100208.
[10] 2020. Kaggle URL: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
[11] Dimpy Varshni1, Kartik Thakral1, Lucky Agarwal1, Rahul Nijhawan2 and Ankush Mittal2, Pneumonia Detection Using CNN based Feature Extraction, 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT) ,DOI: 10.1109/ICECCT.2019.8869364, Coimbatore, India.