Applications of Encoding Methods for False Alarm Elimination in Physiological Monitors: A Comparative Analysis

Sree Sanjanaa Bose S
Department of Biomedical Engineering
KPR Institute of Engineering and Technology,
Coimbatore, India.

Sree Niranjanaa Bose S
Department of Biomedical Engineering
KPR Institute of Engineering and Technology,
Coimbatore, India.

Abstract—**Physiological Monitoring Systems (PMS) are widely evolving to deliver better patient outcomes. In specific, the elimination of false alarms in patient monitors is essential for clinicians to patients who needs attention. In this work, we explore how patient specific variations can be minimized by expressing the input feature vectors in terms of patient independent basis vectors. We used Locality-constrained Linear Coding (LLC) and Fisher Vector Encoding (FVE) for mapping the input features to a patient independent feature space. We obtained a sensitivity of 97.54%, specificity of 96.93%, and overall classification accuracy of 97.41% with LLC-linSVM, whereas with FVE-linSVM we obtained a sensitivity of 98.06%, specificity of 97.62%, and overall classification accuracy of 97.96%. Further we explored the concept of fusing the LLC and FVE mapped features to emulate, looking at the original feature vectors from two perspectives in an effort to improve the performance. We obtained an overall classification accuracy of 98.16%, sensitivity of 98.55%, and specificity of 97.84% with LLC-FVE-linSVM which achieved a performance improvement of 1.01% absolute for classification accuracy, 3.15% absolute for sensitivity and 0.60% absolute for specificity with respect to the baseline system. The multiparameter intelligent monitoring in intensive care II (MIMIC-II) dataset were used for the experiments performed in this work.**

Keywords— PMS, Patient Monitors, SVM, LLC, FVE, MIMIC II Database

I. INTRODUCTION

Continuous vital signs monitoring using physiological monitors has been evolving to deliver better patient outcomes. The physiological monitors also known as Patient monitoring systems (PMS) monitors the human vital signs such as blood pressure (BP), respiration rate (RR), oxygen saturation (SpO2), and heart rate (HR). These vital signs help to identify the condition of the patient. The survey on alarms generated from the PMSs installed in ICUs demonstrate that 88% to 99% of them were false alarms and of less clinical importance. [1]. The oisy environment created in the hospital especially in the ICU due to the myriad of alarms affects the working condition of health workers and thereby deterioration in patient health. The temporary solution adopted by the medical clinicians to this problem is to mute the alarm provision in the sophisticated monitoring device which has resulted in patient death. [2]. The conventional method that was employed in most of the PMSs installed in the ICUs is Early Warning Scores (EWS)

system. The EWS scoring system adopts a methodology in which every critical physiological vital parameter is assigned a numerical score. The outcome of the EWS system is the result of aggregated score and the alarm triggers when the sum exceeds the stable threshold. The ignorance of the relation between the vital parameters and the association of high error-rate with manual scoring are the disadvantages of the EWS system [3]. Therefore the principal focus of the research is the substantial reduction of alarms due to error and maintaining a stable and high degree of sensitivity. This can be achieved through training the machine learning algorithms by feeding the data collected from bed-side monitors to detect the condition of the patients. Machine learning was developed with the aim of building systems that is adaptable to environment and learning from experience. The machine learning algorithms are widely used for different applications.

Development in the area of machine learning has a great impact in health care. Support vector machine (SVM) is one of the popular methods that belongs to the class of supervised learning algorithms [4]. It is widely used for classification and regression in various applications [5]. SVM can be applied for accurate classification of normal and abnormal patients [6]. The classification performance can be enhanced to a significant extent by adopting certain techniques selection of kernels that matches the feature vectors, choosing of feature vectors that perfectly fits the kernels or feature vector linearization that maps the given vector to a higher dimensional space followed by fitting to a linear kernel [7].

The four physiological vital signs from the PMS has been employed as inputs for the SVM classifier. The kernel trick is used for the experimenting the performance of the SVM classifier. The three different kernels used are polynomial, linear and radial basis function (RBF).On comparing the performance of RBF kernel with other kernels, the RBF-linSVM was deployed. But due to the computational complexity and inefficiency of the non-linear kernels being high, the linear SVM classifier is preferred [8]. To linearize the inputs the feature mapping techniques are explored thereby increasing the performance of PMSs.

The process of representing an input data sequence by the weighted linear summation of codebooks or basis vectors,

generated from the feature mapping techniques is known Sparse Coding (SC). After transformation, the SC fails to preserve the locality constraints which is a major disadvantage [9]. In SC, the location of features in transformed feature space will not be as close as situated in the original feature space. Yu et al [10] enunciated the importance of presence of locality should be higher compared to the sparsity. Because of the drawback of SC and to get better results than the SC, local coordinate coding (LCC) was proposed. But, SC and LCC had similar levels of complexity in computation. To overcome the problems, the locality-constrained linear coding (LLC) was suggested due to its computational efficiency and reduced time for implementation. Similar to SC, the input features are represented as linear combination of basis vectors in LLC, but the additional factor is the location constraint [11]. LLC performs the function of mapping the features to a higher dimensional space in order to linearize them. On comparing the performance of SC with LLC, the latter effectively enhanced the classification of SVM [12]-[14].

Another feature mapping techniques used for transforming features to higher dimensionality is the Fisher vector encoding(FVE). The dimension of Fisher vector encoding is 2KD , where D denotes the feature vector dimension and K represents the Gaussian number. However, FVE can compute the feature vectors for the higher dimensionality with the minimum number of Gaussians.

In this work, to enhance the performance of PMS, the use of feature mapping techniques LLC and FVE were explored. Because of the computational inefficiency of non-linear kernel, we use the linear classifier. The inputs (HR, RR, SpO2 and RR) of PMS are non-linear in nature and it is patient dependent. These vital signs should be linearized before classification through lin-SVM classifier. The physiological signs which are used as inputs are linearized using LLC and FVE and then classified using linear SVM backend classified. The LLC- linSVM and FVE-linSVM backend classifiers were experimented using the inputs. The proposed approach make use of the best performing feature mapping methods, by fusing the features of FVE-linSVM and LLC-linSVM. It was found that a noteworthy performance improvement was achieved by FVE-linSVM and the feature fused PMS system when compared to baseline PMS.

II. SYSTEM DESCRIPTION

A. MIMIC – II Database

The multi-parameter intelligent monitoring in intensive care II (MIMIC II) dataset is used in this work for all the experiments. It is a publicly and freely available ICU dataset which consist of wide range of data collected from a very large population of ICU patient. The dataset comprehends two major components of data namely, physiological waveforms and clinical data. It comprises of 125Hz high temporal resolution physiological waveforms and time derived series of HR, RR, BP and SpO2, and alarms generated by the health monitors. The four vital parameters used in the work are associated with a class

label 0 or 1 which indicates the normal and abnormal health condition respectively [15].

B. Locality - constrained Linear Coding (LLC)

Locality-constrained linear coding (LLC) was proposed primarily for image classification by Wang et al [11]. LLC replaces the kernel trick by mapping the input vectors to a higher dimensionality explicitly and then classify them using linear SVM backend classifier. In coding schemes, each input data sequence is denoted by a weighted linear combination of k-nearest codewords from the set of basis vectors or codebook vectors. For each input, the locality constraint is used for selection of k-nearest neighbours. The Lindo Buzo Gray (LBG) algorithm is identical to K-means clustering. The codebook vectors or basis vectors is the result of collection of centroids of each clusters.

Assume the input feature vector as $S=[x_1, x_2 \dots x_n]$ and the codebook generated by the LBG algorithm as $B=[b_1, b_2 \dots b_n]$. The following represents the optimization criterion for LLC

$$\min_c \sum_{i=1} \left\| x_i - B\mu_i \right\|^2 + \lambda \left\| d_i \square \mu_i \right\|^2 \qquad (1)$$

such that $1^T\mu_i=1$, $\forall i$ where μ_i and d_i depict the codeword for each input feature vector x_i and locality adaptor respectively. The element-wise multiplication is denoted by $o\ dot$ operator. The locality adaptor provides the degree of freedom for each basis vector proportional to its corresponding feature vector x_i.

$$d_i = \exp(\frac{dist(x_i, B)}{\sigma}) \qquad (2)$$

The dist(x,y) is a column matrix which contains the Euclidean distance measured between xi and each basis vector. The locality adaptor ensures that similar input feature vectors have similar representations in the transformed higher dimensional feature vector space. σ denotes the weights adjustable decay speed for locality adaptor. By calculating the difference between max(dist(x,y)) and dist(x,y), the xi parameter can be normalized to (0,1). The constraint 1T μi=1 is known as the shift invariant component of LLC. The presence of non-zero values shows that any nearest neighbours are used for the encoding process replacing the need for utilizing the complete codebook. This method is called the fast approximation LLC which is denoted as

$$\min_c \sum_{i=1}^{N} \left\| x_i - Bc_i \right\|^2 \qquad (3)$$

C. Fisher Vector Encoding

FVE was introduced by Perronin et al [16] for image classification having larger number of data. It involves the computation of first and second order derivatives between the Gaussian centres and feature vectors. FVE depends on the Gaussian Mixture model (GMM) whose parameters are matched to the feature vectors. Following this, the derivatives of the probability of the model are encoded with respect to these parameters. The computation of

codebooks required for the Fisher vector encoding are generated with the help of Gaussian mixture models (GMM) [12]. The GMM uses expectation maximization (EM) algorithm which does soft assignments of input features. It is a parametric probability density function. Based on the posterior probabilities, the mixture components are allocated with appropriate input features. Assume the input features as $x = [x_1, x_2 .... x_n]$ and q_{ki} as the soft assignments to the N input features to the K gaussians. It can be represented as,

$$q_{ki} = \frac{p(x_i \mid \mu_k, \Sigma_k)\prod_k}{\sum\limits_{j=1}^{K} p(x_i \mid \mu_j, \Sigma_j)\prod_j}, K = 1, 2, ... k \qquad (4)$$

The mean u_k and the covariance v_k derivative features are computed for each K=1,2,... k.

$$u_k = \frac{1}{N\sqrt{\prod_k}} \sum_{i=1}^{N} q_{ik} \Sigma_k^{-1/2} (x_t - \mu_k) \qquad (5)$$

$$v_{k=} \frac{1}{N\sqrt{2\prod_k}} \sum_{i=1}^{N} q_{ik}[(x_t - \mu_k)\Sigma_k^{-1}(x_t - \mu_k) - 1] \qquad (6)$$

The GMM parameters μk, $\prod_k$ and $\sum_k$ symbolize the mean, prior probabilities and covariance matrix of each Gaussian component respectively. The algorithm assumes the covariance matrices Σ_k to be diagonal. This reduces the computation time of the feature vectors. The FVE features of the input vectors are represented by concatenating uk and vk for all K Gaussian centers.

$$f_{fisher} = [u_1^T, v_1^T, ... u_K^T, v_K^T]^T \qquad (7)$$

III. EXPERIMENTS AND RESULTS

The MIMIC II database is a freely available data which contains vital signs data collected from 420 patients in the ICU during their treatment stay at hospital. [16]. Among the 420 patient data, the four vital signs of 19 patients was missing and they were excluded from the study. The 401 patient data was divided as training and testing data. The training set contained 300 patient data and the testing set included 101 patient data. The training data was clubbed into seven subsets with each training subset containing 50,000 samples and with each testing set containing 20,000 samples. Seven trials were conducted by pairing the testing and training sets. By calculating the average of seven trials, the classification accuracy was computed. The evaluation of performance of the classifier was based on the sensitivity, specificity and overall accuracy of the system. LIBSVM was deployed for all the experiments.

A. Baseline System

The baseline system was developed with four vital signs as inputs to the SVM backend classifier. The classifier was tested with three different kernels such as polynomial, linear and Radial Basis Function (RBF).

TABLE I

RESULTS OF SVM WITH DIFFERENT KERNELS (IN %)

Kernel	Accuracy	Sensitivity	Specificity
Linear	76.95	76.84	93.85
Polynomial	87.95	86.72	92.83
Radial Basis Function	97.17	95.54	97.25

Table I lists the performance of the baseline system using different kernels. It clearly depicts the RBF kernel outperforms the linear kernel. Ying [18] have found that deterioration in alarm and no-alarm accuracy has been the cause of several life-threatening conditions. RBF-SVM is considered as the baseline PMS.

B. LLC-LinSVM system

In LLC, the codebooks make use of the training data for generating the LLC features. Both the training and testing data uses the same codebook generated with training data. Table II shows the results obtained by LLC-linSVM PMS for different codebook sizes of 256, 512, 1024 and 2048.

TABLE II

RESULTS OF LLC-LINSVM SYSTEM USING DIFFERENT CODEBOOK SIZES (IN %)

Codebook Size	KNN	Accuracy	Sensitivity	Specificity
256	20	96.04	88.62	98.29
512	20	96.73	90.08	98.74
1024	20	97.17	91.23	98.97
2048	20	97.41	97.54	96.93

The LLC-linSVM system performance has improved significantly compared to PMSs using RBF SVM backend classifier (RBF-SVM). dimension of the FVE features with 256 cluster size was 2048. The dimension size was same as the LLC features.

C. FVE-linSVM system

FVE experiment were performed by increasing the number of clusters from 8 to 512. Table III lists the results of FVE- linSVM system obtained using different cluster size. It was observed that 256 cluster size gave the best performance. The dimension size was same as the LLC features

TABLE III

RESULTS OF FVE-LIN SVM SYSTEM USING DIFFERENT CLUSTER SIZES (IN %)

Cluster Size	Accuracy	Sensitivity	Specificity
8	95.72	87.98	98.06
16	96.75	90.35	98.68
32	97.39	92.01	99.02
64	97.62	92.65	99.12
128	97.66	92.65	99.12
256	97.96	98.06	97.62
512	97.71	92.70	99.23

From Table III, it is seen that FVE-linSVM system performs better for cluster size 256 and as the cluster size increases there is a decrease in accuracy and sensitivity.
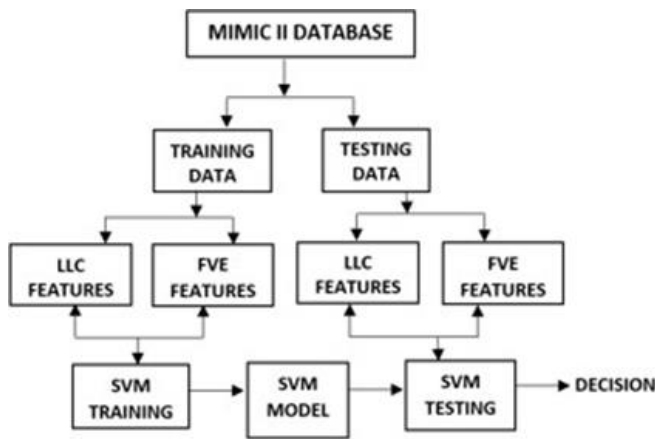
Fig. 1 Block Diagram of Proposed System

D. LLC-FVE-linSVM system

Feature level fusion of the mapped features were performed to improve the performance of the PMSs. By combining the vital sign data at the feature level more information about the input feature is obtained. The features obtained from the LLC mapping technique and the features obtained from the FVE mapping technique were concatenated. The training and testing data with the fused features were given as its input to the linear SVM classifier as shown in Fig 1. The performance of the system was visualized using confusion matrix or contingency matrix.

TABLE IV
PERFORMANCE COMPARISON OF DIFFERENT SYSTEMS (IN %)

System	Accuracy	Sensitivity	Specificity
RBF-SVM	97.17	95.54	97.25
LLC-linSVM	97.41	97.54	96.93
FVE-linSVM	97.96	98.06	97.62
LLC-FVE-linSVM	**98.16**	**98.55**	**97.84**

IV. DISCUSSION

The baseline system that uses the vital signs directly as features achieved an accuracy of 97.17%, sensitivity of 95.54%, and specificity of 97.25% from RBF-SVM based PMS. In addition, we explored how patient specific variations can be minimized by expressing the input feature vectors in terms of patient independent basis vectors. We used Locality-constrained Linear Coding(LLC) and Fisher Vector Encoding(FVE) for mapping the input features to a patient independent feature space. These features are linearized and therefore the Linear SVM was used for decision making. Th LLC-linSVM resulted in the sensitivity of 97.54%, specificity of 96.93%, and overall classification accuracy of 97.41%. It achieved a performance improvement of 0.24% absolute for classification accuracy, 2.09% absolute for sensitivity with respect to the baseline system. But there was a deterioration of 0.33% absolute for sensitivity. Whereas with FVE-linSVM, we obtained a sensitivity of 98.06%, specificity of 97.62%, and overall classification accuracy of 97.96% which achieved a performance improvement of 0.81% absolute for classification accuracy, 2.63% absolute for sensitivity and 0.38% absolute for specificity with respect to the baseline system.

To combine the linearization effects of both LLC and FVE, we experimented using both mapped features to train the Linear SVM. From Table IV, it is observed that the LLC-FVE-linSVM backend classifier has achieved a performance improvement of 1.01% absolute for classification accuracy, 0.30% absolute for sensitivity and 0.60% absolute for specificity compared to the RBF-SVM baseline system. The codebook size of 256 gave optimal results for the tested dataset. The optimal codebook size is essential to capture the useful information from data and also ensure that it doesn't occupy more space when it is deployed in embedded systems.

We explored the Principle Component analysis (PCA) and Linear Discriminant Analysis (LDA) in an effort to decrease the number of features thereby enhancing the performance of PMS. But the results obtained were not satisfactory and did not contribute in improving the PMS performance. The future work will focus on how effectively these techniques can be used for better improvement.

V. CONCLUSION

In this work, the feature mapping techniques have been explored through locality constrained linear coding (LLC) and Fisher Vector encoding (FVE). These techniques transformed the input features to a patient independent feature space. These features are classified with a linear (SVM) backend classifier and the performance of multi-parameter patient monitors (PMSs) was improved. In the performance comparison of our proposed system with the baseline PMS employing RBF kernel, the LLC-FVE-linSVM PMS achieved a significant performance improvement in overall classification accuracy, sensitivity and specificity of 1.01%, 3.15% and 0.60% absolute respectively. The future work will focus on how other encoding techniques can be used effectively techniques can be used for better improvement of PMS system.

REFERENCES

[1] F. Schmid, M. S. Goepfert, D. Kuhnt, V. Eichhorn, S. Diedrichs, H. Reichenspurner, D. A. Reuter, The wolf is crying in the operating room: patient monitor and anesthesia workstation alarming patterns during cardiac surgery, Anesthesia & Analgesia, vol. 112, no. 1, pp. 78-83, 2011.

[2] F. B. Rejab, K. Nouira, and A. Trabelsi, Real Time SVM for Health Monitoring System, in Brain Informatics and Health, Springer, pp. 301- 312, 2014.

[3] L. Tarassenko, D. A. Clifton, M. R. Pinsky, M. T. Hravnak, J. R. Woods, and P. J. Watkinson, Centile-based early warning scores derived from statistical distributions of vital signs, Resuscitation, vol. 82, no. 8, pp. 1013-1018, Aug. 2011.

[4] V. Vapnik, Statistical learning theory, Vol. 1, New York: Wiley, 1998.

[5] A. Sumesh, Rameshkumar, K., Mohandas, K., and R. Babu, S., Use of Machine Learning Algorithms for Weld Quality Monitoring using Acoustic Signature, Procedia Computer Science, vol. 50, pp. 316 - 322, 2015.

[6] S. Khalid, D. A. Clifton, L. Clifton, and L. Tarassenko, A Two-Class Approach to the Detection of Physiological Deterioration in Patient Vital Signs, With Clinical Label Refinement, in IEEE Transanction on information Technology in Biomedicine, vol. 16, no. 6, pp. 1231-1238, 2012.

[7] A. Vedaldi and A. Zisserman, Efficient additive kernels via explicit feature maps, in IEEE Transaction on Pattern Analysis and Machine Intelligence, vol 34, no. 3, pp.480-492, 2012.

[8] S. Maji, A. Berg, J. Malik, Classification using intersection kernel support vector machines is efficient, in Computer Society Conference on Computer Vision and Pattern Recognition, pp. 1-8, June 24-26, 2008.

[9]   J. Yang, K. Yu, Y. Gong and T. Huang,Linear spatial pyramid matching using sparse coding for image classification, in IEEE conference on Computer Vision and Pattern Recognition, pp. 1794-1801, 2009.

[10]   K. Yu, T. Zhang and Y. Gong, Nonlinear learning using local coordinate coding, Advances in neural information processing systems, pp. 2223- 2231, 2009.

[11]   J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, Locality-constrained linear coding for image classification, in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3360-3367, 2010.

[12]   K.T. Sreekumar, R. Gopinath, M. Pushparajan, A.S. Raghunath,C.S. Kumar, K.I. Ramachandran and M. Saimurugan, Locality constrained linear coding for fault diagnosis of rotating machines using vibration analysis, in Annual IEEE India Conference (INDICON), pp. 1-6, 2015.

[13]   R. Gopinath, C. Santhosh Kumar, and K. I. Ramachandran. "Fisher vector encoding for improving the performance of fault diagnosis in a synchronous generator.", in Measurement, vol. 111,pp. 264-270, 2017

[14]   Y. Liang, M. Song, J. Bu, and C. Chen, Colorization for gray scale facial image by locality-constrained linear coding, in Journal of Signal Processing Systems, vol. 74, no. 1, pp. 59-67, 2014.

[15]   J. Lee, D. J. Scott, M. Villarroel, G. D. Clifford, M. Saeed, and R.G. Mark, Open-access MIMIC-II database for intensive care research, in International Conference of IEEE in Engineering in Medicine and Biology Society (EMBC), pp. 8315-8318, 2011.

[16]   F. Perronnin, , J. Sanchez, and T. Mensink, Improving the fisher kernel for large-scale image classification, In Computer Vision ECCV 2010, Springer,Berlin,Heidelberg, pp. 143-156. 2010.

[17]   MIMIC   II   Database, [Online]. Available: www.physionet.org/physiobank/database/mimic2db.

[18]   Chang, Chih-Chumg and Lin, Chih-Jen, LIBSVM: a library for support vector machines, in ACM Transactions on Intelligent Systems and Technology (TIST), vol.2, no. 3, pp.27, 2011.

[19]   Ying Zhang: Real-Time Development of Patient-Specific Alarm Algorithms for Critical Care, in 29th International Conference of IEEE in Engineering in Medicine and Biology Society (EMBC), pp.4351-4354, 2007.

[20]   Chatfield, Ken, S. Victor Lempitsky, Andrea Vedaldi, and Andrew Zisserman, " The devil is in the details: an evaluation of recent feature encoding methods." In BMVC, vol. 2, no. 4, pp. 8. 2011.