ONLINE KERNEL AMGLVQ FOR ARRHYTHMIA HEARBEATS CLASSIFICATION

[a]**Elly Matul Imah,** [b]**R. Sulaiman**
[a,b]Mathematics Department, Universitas Negeri Surabaya
Kampus Unesa Ketintang, Surabaya, Indonesia
Email: ellymatul@unesa.ac.id

Abstract

This study proposes Online Kernel Adaptive Multilayer Generalized Learning Vector Quantization (KAMGLVQ) for handling imbalanced data sets. KAMGLVQ is extended version of AMGLVQ that used kernel function to handling non-linear classification problems. Basically AMGLVQ is vector quantization based learning. The vector quantization based learning is very simple algorithm that can be applied to the multiclass problem and the complexity of LVQ can be controlled during training process. KAMGLVQ works at online kernel learning system that integrating feature extraction and classification. The architecture network of KAMGLVQ consists of three layers, input layer, hidden layer, and an output layer. The hidden layer of KAMGLVQ is adaptive; this algorithm will generate a number of hidden layer nodes. The algorithm implement on real ECG signals from the MIT-BIH arrhythmias database and synthetic data. The experiments showed that KAMGLVQ able improve the accuracy of classification better than SVM or back-propagation NN; also able to reduce the time computational cost.

Keywords: KAMGLVQ, SVM, Backpropagation, arrhythmia, ECG, imbalanced data set

INTRODUCTION

The imbalanced data set is a classification problem because almost learning algorithm is designed for the balanced dataset. The imbalanced dataset is a set of training data wich one of the classes is represented by small number than the other class. There are two approaches for dealing this problem, with the internal and external approach. Internal approach is handling imbalanced data set by using algorithm modification [1], [2]. External approach is handling it by using resampling: undersampling and oversampling [1], [3]. Recently some researchers use combining of internal and external approaches [4]. In our study, the proposed Online Kernel Adaptive Generalized learning Vector Quantization (KAMGLVQ) to handle the imbalanced dataset by using combining external approach and internal approach.

Online KAMGLVQ is non-linear version of Adaptive Multilayer Generalized Learning vector Quantization (AMGLVQ) that has been proposed in our previous work [5]. KAMGLVQ is integrating feature extraction and classification and work at online learning system. Online learning is commonly used techniques for training in the dynamic condition of data. It is successfully used in the neural network. KAMGLVQ is developing from Generalized Learning Vector Quantization (GLVQ) that proposed by A. Sato [6]. GLVQ is a part of Vector Quantization based learning.

The Learning Vector Quantization (LVQ) classification is learning algorithm that based on a distance measure, which quantifies the similarity of given data with so-called prototypes weight or codebook vectors representing the classes. The prototypes are determined from training dataset in a training process and interpreted in a straightforward way that they capture essential feature of the data in the same space. The complexity of LVQ can be controlled during training process and it is a very simple algorithm that can be applied to the multiclass problem. Some researcher mixed of LVQ with the other quantization based learning, like SOM or Neural Gas that including the neighborhood operation in the learning process [6–9].

KAMGLVQ is extended version of AMGLVQ that used kernel function to handling non-linear classification problems. Basically AMGLVQ is vector quantization based learning. AMGLVQ have been handling an imbalanced dataset, inconsistency between feature extraction and classification problem, but in some case that non-linear, the boundary class that has been created is not smooth. The algorithm was evaluated on real ECG signals from the MIT-BIH arrhythmia database, iris dataset, and some synthetic data generations.

The aim of our study is to develop ECG smart device. Our previous study in this field have been published on [10–12]. ECG smart device is a part of automatic heart beats classification. It has attracted much interest for research because it can save cardiologist from looking for arrhythmia beats in a sheer amount of ECG heartbeats data. Such automatic classification will enable cardiologist to diagnose cardiovascular disease (CVD) timely and to make home care possible. Many algorithms have been proposed to automatically classify life-threatening arrhythmia beat types from ECG data. They are focus on feature extraction [13], [14] and some of researcher focus on classification process [15], [16]. Weighted support vector machine (wSVM) classifier [16] and weighted conditional random fields [17] have been proposed in the fields of heartbeat classification to overcome the class imbalanced. However, neither of them tackles the unclassifiable heartbeat and inconsistency problem between feature extraction and classification.

MIT-BIH Arrhythmia database often used for automatic heartbeats classification studies. The has characteristics strongly imbalanced classes [17], and nonlinearly separable. So, in this study we used the MIT-BIH database to evaluate our proposed method. Conventional learning methods generally perform poorly because they are designed for balance data, so they cannot solve some imbalanced classes' data, without preprocessing to balancing the dataset. A set of training data is said to be imbalanced if one of the classes is represented by a small number of cases compare to the other classes.

The paper organization is as follows. Section II describes system overview and section III describes a pre-processing technique to reduce the ECG noise signal that caused of wandering the isoelectric by using cubic spline and then

extract a signal in beat basis. Section fourth describe the proposed method, online KAMGLVQ then the experiment result and discuss describe on section IV. The experiment result and discuss describe on section V and the last section is section VI that describes the conclusions of this study.

OVERVIEW OF SMART CARDIO DEVICE SYSTEM

The overall system of Smart Cardio Device is shown in Figure 1. The device was divided into three main components, the first component is a circuit used to multiply and filter ECG signal. It was designed to make the ECG signals high so that those signals could be executed on Beagle board. The second component is used to run classification process. The third component is a mobile device as interfacing of our system. The study, focus on the second component.

The detail of system overview is shown in Figure 2. In this system, we built ECG module then executed the signal on Beagle board. We used Android mobile device for interfacing the signal and LCD touch-screen as user input interaction. To connecting the signal from ECG module to mobile device we used wirelessly. From Android mobile device, the user can show the ECG signal and enter the input data to interaction to the system. From the overall system of the smart cardio device, this study focuses on classification algorithm.
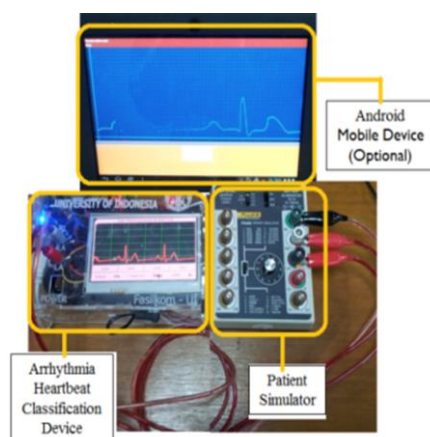

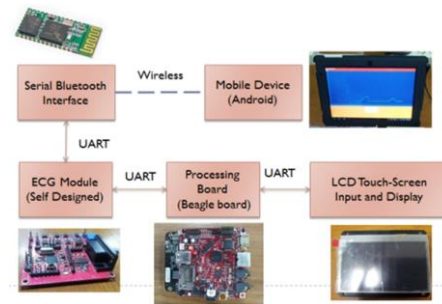Figure 1. Overall System of
Smart Cardio Device

Figure 2 . System overview of
Smart Cardio Device

SIGNAL PREPROCESSING

This study use MIT-BIH arrhythmia data sets, it's available on Physionet database [18]. MIT-BIH Arrhythmia database contains 48 records, which each record contains a pair of 30-minutes ECG signal. The frequency of the ECG data was 360 Hz. In this study, we only used MLII lead to detection the arrhythmias heartbeat. One of the most importance parts to detect of arrhythmia heartbeats is arrhythmia heartbeat annotation. KAMGLVQ implemented to classify the arrhythmia heartbeats annotation. The classes of arrhythmia heartbeats annotation that we used in this study are Normal Beat (NOR), Left Bundle Branch Block beat (LBBB), Right Bundle Branch Block beat (RBBB), Premature Ventricular Contraction beat(PVC), Paced beat (P), Atrial Premature beat (AP), Fusion of Ventricular and Normal beat(fVN), Fusion of Paced and Normal beat(fPN), Nodal (Junctional) Escape beat (NE), Aberrated Atrial Premature beat (aAP), Ventricular Escape (VE), Nodal (junctional) Premature beat (NP), Arial Escape beat (AE) and Supraventricular Premature beat (SP).

At the first, the continuous ECG signals will be transformed into individual ECG beats that would be given into the classifier to label classes. To do so, series of ECG signals passed two preprocessing steps starting from the removal of baseline noise and followed by beat extraction to extract the individual beats. Those entire two steps will be described in detail in the following section.
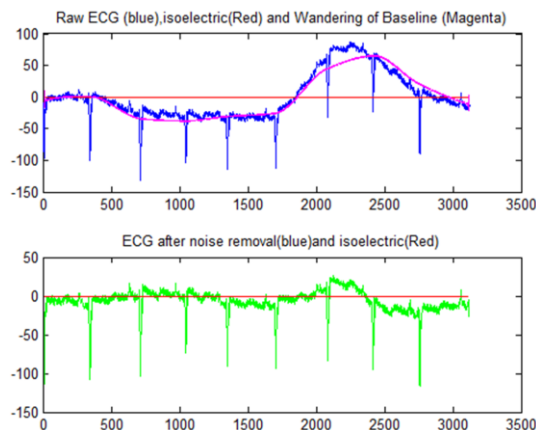
Figure 3. Baseline estimation and removal with cubic splines interpolation

Baseline Noise Removal

Baseline noise removal is the first step of preprocessing phase because the baseline noise or baseline wander is a condition where the to generate ECG signal shifted up or down from isoelectric. The noise or baseline wandering is caused by low-frequency activity during the ECG recording. It is disturb and able to made a misinterpretation of ECG signal analysis.

We used cubic spline filtering as developed by Badilini [19]. The study used cubic spline interpolation to reduce the baseline noise signal, because this interpolation can remove the wondering of isoelectric with no significant influent to the ECG signal. The spline is generated form ECG PR-segment, and then beaseline noise is reduced by substracting the estimated cubic splinefunction to the raw data. The detil can be seen on Figure 3.

Beat Extraction

After noise wander of baseline was removed, the ECG signal will be extracted into single beat in this phase. The extraction of ECG signal to single beat using approximation the width of individual beat into 300 sample points. Sample point of the single beat is centered on R peak. The beat was extracted by cutting of the continuous signal at every R peak, start from R-150 to R+149. Hence the extraction result is a single beat that contain 300 sample points. See Figure 4 for the illustration of beat extraction.
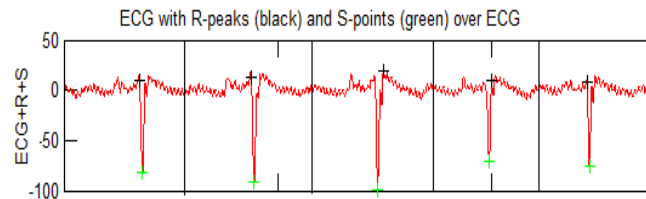

Figure 4. The cutoff technique used in this transformation process

ONLINE KERNEL ADAPTIVE MULTILAYER GENERALIZED LEARNING VECTOR QUANTIZATION (KAMGLVQ)

Online KAMGLVQ is a non-linear model of AMGLVQ that proposed in the previous study [5]. This algorithm used the integrating model of feature extraction and classification. Feature extraction and classification are related each other. Good feature extraction can result in the good performance of classification. This study used an integrated model of feature extraction and classification, because with the integrated system, the inconsistency between feature extraction and classification able to minimize. Our methods are a difference with the conventional classification system model, the conventional classification system model does not integrate the feature extraction phase and a classification phase. System block diagram of the conventional classification system and integrated system can be seen in Figure 5 and 6.
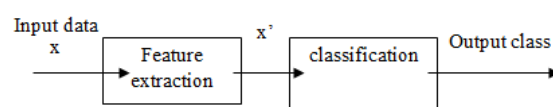

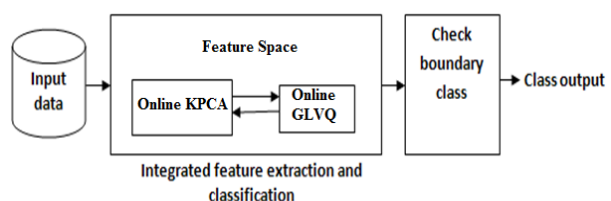Figure 5. Block diagram of conventional classification system

Figure 6. Block diagram of integrated feature extraction and classification
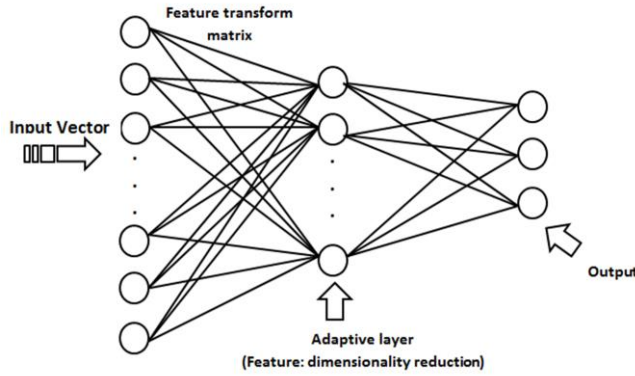
Figure 7. Online KAMGLVQ Architecture

The algorithm is proposed for a nonlinear system of classification. The algorithm has three layers on architecture neural network, the first layer is input layer, then the second layer is a hidden layer that has an adaptive number of neuron and the third layer is output layer, for the detail sees in Figure 4.

The architecture KAMGLVQ on Figure 7, show that feature reduction step is on the hidden layer. The number of nodes in hidden layer depends on a number of features that has been reduced by this algorithm. From input layer, a full feature of data will be transformed to the feature space and then reduces the dimension of data. KAMGLVQ also adaptive like AMGLVQ, it is caused the number of node in a hidden layer is adaptive depend on the number of features and the number of features depends on matrix transformation that is composed of selected data. The online KAMGLVQ has two phases, the first phase is optimization of integrating feature extraction-classification and handling imbalanced data set, and then the second phase is handling unknown class.

The first phase is the optimization of integrating feature extraction-classification and handling imbalanced data. The first step, selected data are used to find the principal component sequentially, so the matrix that are used to find the principal component increase by adding the selected input vector that has big variance. The selecting process also able to handle imbalanced data set also. Online KAMGLVQ used algorithms and data level approach for handling imbalanced data sets. The data level approach for handling imbalanced data set is under-sampling, the sample data will be selected by the evaluation measure of classification.

The first step in this phase is data selection to build the matrix transformation that consists of the principal component. The Matrix will be updated with adding the input vector that has the misclassification error positive. If $X_1, X_2, \ldots, X_N \in \mathbb{R}^N$ and $Y_1, Y_2, \ldots, Y_M \in \mathbb{R}^N$ is data train from differences group, X_N from t-th group training data, and Y_N is (t+1)-th group training data. Z is defined new matrix that is used to find the principal component in the feature space, $Z = X \cup Y$, so we can write Z as $Z = \{X_1, X_2, , \ldots, X_N, Y_1, Y_2, \ldots, Y_M\} \in \mathbb{R}^N$.

v_x and v_y span by $\phi(X_1), \phi(X_2), \ldots, \phi(X_N)$ and $\phi(Y_1), \phi(Y_2), \ldots, \phi(Y_N)$, then v_x dan v_y can define as linear combination of α and $\phi(X_i)$ for v_x, ω and $\phi(Y_i)$ for v_y. We can also write as $v_x = \sum_{i=1}^{N} \alpha_i \phi(X_i)$ and $v_y = \sum_{i=1}^{N} \omega_i \phi(Y_i)$.

Matrix gram kernel of $X_1, X_2, X_2, \ldots, X_N \in \mathbb{R}^N$ is defined in Equation (1):

$$K^{XX} = \begin{bmatrix} \phi(X_1)\phi(X_1) & \ldots & \phi(X_1)\phi(X_N) \\ \vdots & \ddots & \vdots \\ \phi(X_N)\phi(X_1) & \ldots & \phi(X_N)\phi(X_N) \end{bmatrix} \quad (1)$$

with the Mercer theorem equation can be written in Equation(2)

$$K^{XX} = \begin{bmatrix} K(X_1, X_1) & \ldots & K(X_1, X_N) \\ \vdots & \ddots & \vdots \\ K(X_N, X_1) & \ldots & K(X_N, X_N) \end{bmatrix} \quad (2)$$

Then matrix Gram kernel of $Y_1, Y_2, Y_2, \ldots, Y_M \in \mathbb{R}^N$ can be defined as Equation (3).

$$K^{YY} = \begin{bmatrix} \phi(Y_1)\phi(Y_1) & \ldots & \phi(Y_1)\phi(Y_M) \\ \vdots & \ddots & \vdots \\ \phi(Y_M)\phi(Y_1) & \ldots & \phi(Y_M)\phi(Y_M) \end{bmatrix} \quad (3)$$

The Equation (3) can be written into Equation (4).

$$K^{YY} = \begin{bmatrix} K(Y_1, Y_1) & \ldots & K(Y_1, Y_M) \\ \vdots & \ddots & \vdots \\ K(Y_M, Y_1) & \ldots & K(Y_M, Y_M) \end{bmatrix} \quad (4)$$

If $Z = X \cup Y$

$$= \{X_1, X_2, , \ldots, X_N, Y_1, Y_2, \ldots, Y_M\} \in \mathbb{R}^N$$

So the covariant matrix of $\phi(Z)$ is shown in Equation (5):

$$K_z = \frac{1}{N+M}\left(\sum_{j=1}^{N+M} \phi(Z_j)\phi(Z_j)^T\right)$$

$$K_z = \frac{1}{N+M}\left(\sum_{j=1}^{N} \phi(X_j)\phi(X_j)^T + \right.$$

$$\left. \sum_{j=1}^{M} \phi(Y_j)\phi(Y_j)^T\right) \qquad (5)$$

The Eigenvalue in the feature space can be written as Equation (6).

$$\gamma u = K_z u \qquad (6)$$

u span of $Z = X \cup Y$ so like an PCA we can write Equation (6) as seen in Equation (7).

$$\gamma K\xi = \frac{1}{N+M} K_z^2 \xi$$

$$(N+M)\gamma\xi = K_z\xi \qquad (7)$$

where

$$K_z = \begin{bmatrix} K^{XX} & K^{XY} \\ K^{YX} & K^{YY} \end{bmatrix}, \ K^{YX} = (K^{XY})^T$$

Then Center the gram matrix K by using Equation (8).

$$K_{i,j} = \left(K - 1_{(N+M)}K - K1_{(N+M)} + \right.$$

$$\left. 1_{(N+M)}K1_{(N+M)}\right)_{i,j} \qquad (8)$$

Solve Equation (7) to find the principal component of Z after K is centered by using Equation (8). We find T as a principal component matrix that is used to transform data from data space to feature eigenspace.

Then to update the prototype of classification, the vector references should be updated. Let x is input data, w_1 the nearest reference vector that belongs to the same class of x, and likewise let w_2 be the nearest reference vector that belongs to a different class from x. Let us consider the relative distance $\mu(x)$ is defined in Equation (9)

$$\mu(x) = \frac{d_1 - d_2}{d_1 + d_2} \qquad (9)$$

Where d_1, d_2 are the distance of x from w_1 and w_2 respectively, $\mu(x) \in [-1, +1]$ and if $\mu(x)$ is negative, x is classified correctly; otherwise, x is classified incorrectly. In order to improve error rates, $\mu(x)$ should decrease for all input vectors. The evaluation measure is to select the sample data, using Equation (9).

Thus, a criterion for learning is formulated as minimizing cost function S is defined in Equation (10).

$$S = \sum_{i=0}^{N} f(\mu(x_i)) \qquad (10)$$

Where N is the number of input vector for training data recently because system runs in online learning system, and $f(\mu)$ is a monotonically increasing function, e.g. the logistic function $f(\mu) = {}^1\!/_{(1 + \exp(x))}$. For sigmoid functions, the classifier only learns from the training samples lying close to the decision boundary which carry most representative information.

After matrix transformation T is updated, and then we have to synchronize the prototype of classification to new T. Then update the references vector by using Equation (11) and Equation (12).

$$w_1 \leftarrow w_1 + \alpha \frac{\delta f}{\delta \varphi} \frac{4d_2}{(d_1 + d_2)^2}\left(K_{x(i)}T_{K(i)} - \right.$$

$$\left. K_{w_1(i)}T_{K(i)}\right) \qquad (11)$$

$$w_2 \leftarrow w_2 - \alpha \frac{\delta f}{\delta \varphi} \frac{4d_1}{(d_1 + d_2)^2}\left(K_{x(i)}T_{K(i)} - \right.$$

$$\left. K_{w_2(i)}T_{K(i)}\right) \qquad (12)$$

The second phase is finding the boundary class for incremental class on testing phase. To get the boundary class we use the principal of the heuristic clustering. We find the threshold of every class to be the boundary of the class using Equation (13).

$$r = \frac{\sum_{i=1}^{N} d_n}{N} + 3\sigma \qquad (13)$$

The testing data that out of boundary class output in the testing phase, will be created a new class as unclassifiable heartbeat and this class is an increment, depend on how far the differences in the data.

RESULT AND DISCUSSION

In this phase we will validate our method for the classification problem, we used synthetic data and Arrhythmia heartbeats classification problem. We would prove that our method was able to apply in the balanced dataset or imbalanced dataset. Other than confusion matrix; we used to recall, precision, true positive rate, false positive rate and F-measure for evaluation measure of the imbalanced dataset.

Evaluation measure

An imbalanced dataset, not only is the class distribution is skewed, the misclassification cost is often uneven too. The minority class example is often more important than the majority class examples[20]. Some relevant evaluation measurements are confusion matrix, recall, precision, and F-measure. The confusion matrix is useful when accessing the performance without taking the cost into consideration. It is used as a basis for various measures, such as precision and recall. Precision is defined in Equation (14) and recall in Equation (15).

$$precision = \frac{TP}{TP+FP} \qquad (14)$$

$$recall = \frac{TP}{TP+FN} \qquad (15)$$

F-measure is a classification evaluation metrics that combines precision and recall, that can be defined as in Equation (16).

$$F - measure = \frac{2 \times Recall \times Precision}{Recall+Precision} \quad (16)$$

The study, we use confusion matrix, recall, precision, F-measure to evaluation measurement.

Generate Synthetic Data Sets

We generate random datasets that consist of three classes. The generated data is imbalanced with the number of data in first class is 500, data in a class second class is 5 and the third class is 100. The experiment result shows that KAMGLVQ shows better performance than the SVM. We used SVM because it's standard benchmark classifier recently. The accuracy of multiclass SVM accuracy is 95.21% with the CPU time 5.8407 seconds and KAMGLVQ is 95.54% with CPU time 0.5581 seconds. The contour of topology data can be shown in Figure 8 and Figure 9.

Arrhythmia heartbeats classification

The detailed experiment uses MIT-BIH the Arrhythmia database as follow; data is imbalanced with the imbalanced ratio is 1:100.
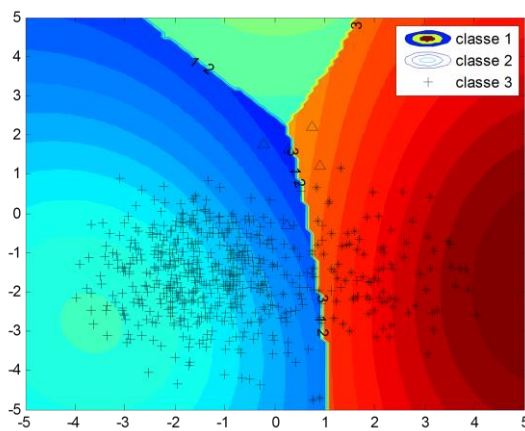

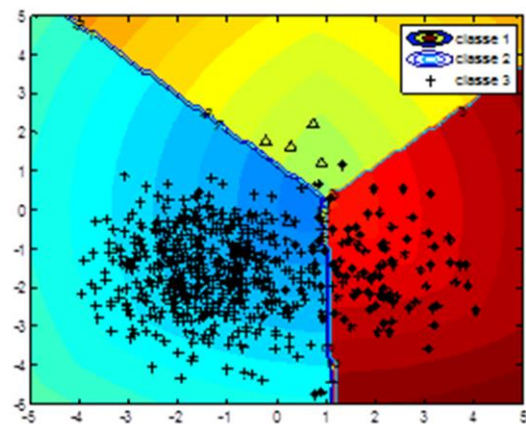Figure 8. The contour topology synthetics data of SVM

Figure 9. The contour topology synthetics data of KAMGLVQ

The class annotation of arrhythmia heart beat are 200 beats for NOR class, 200 beats for LBBB class, 200 beats for RBBB class, 200 beats for PVC class, 200 beats for AP class, 200 beats for fVN class, 200 beats for fPN class, 200 beats for NE class, 138 beats for aAP class, 106 beats for VE class, 47 beats for NP class, 15 beats fo AE class, 2 beats for SP. Every single beat consist of 300 points. Then for the predecessor of our proposed method, we use PCA as feature extraction combining with SVM, back-propagation, GLVQ, and LVQ. The experiment result is shown at Figure10.

The best number of principal component in conventional PCA feature extraction in this staudy is 60 principal component. This principal component we used to reduce the full feature of ECG data signal, the 60 feature as result of feature extraction are used as input data to classify the Arrhythmias heartbeat by using LVQ, GLVQ, SVM, and back-propagation-NN. Then the result we compare with KAMGLVQ the method that we proposed. Accuracy in scenario experiment without unkown class as follow, the accuracy of LVQ is 85.18%, GLVQ 87.94%, SVM 92.4%, Back-propagation-NN 93.54%, and KAMGLVQ is 94.58%. The accuracy with adding unkown class into testing data of LVQ decrease from 85.18% to 80.01%, GLVQ decrease from 87.94% to 82.4%, SVM decrease from 92.4% to 87.01%.
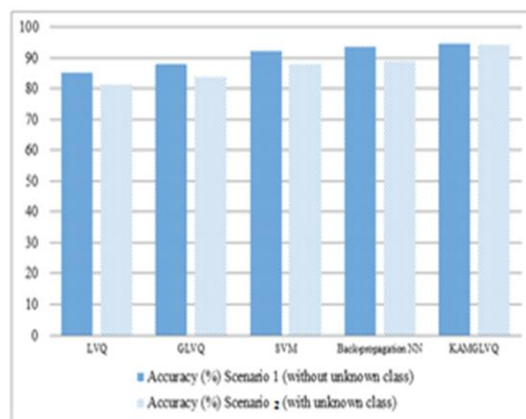

Figure 10. Experiment Result Arrhythmia heartbeats classification without unknown category

Back-propagation-NN decrease from 93.54% to 88.9%, but KAMGLVQ is still show a beter better result than the predecessor, the accuracy of KAMGLVQ in this scenario is 94.5%, decrease only 0.08%. The detail of evaluation measure of KAMGLVQ is shown in Table 1. The precision, Recall, and F-measure of 13 classes Arrhythmia heartbeats show good performance, the average Precision of KAMGLVQ is 0.945, Recall is 0.923, and F-measurement is 0.926. This result shows that KAMGLVQ able to recognize every classes even in a strongly imbalanced dataset.

CONCLUSION

The study shows that KAMGLVQ has a good performance to implementation on imbalanced data sets and show good performance for imbalanced data sets than the SVM and Backpropagation. Beside of that KAMGLVQ is able to generate new class if the testing phase detected a new type of heartbeat. The difference with the others conventional classification, KAMGLVQ is able to generate the new class, because this algorithm adopted concept boundary class of heuristic clustering.

ACKNOWLEDGEMENT

Table 1. Evaluation Measure Of Arrhythmia Data Set

Class	Precision	Recall	F-measure
'RBBB'	0.99	0.961	0.975
'NOR'	1	0.833	0.909
'LBBB'	1	0.971	0.985
'AP'	1	1	1
'fVN'	1	1	1
'PVC'	0.960	0.951	0.955
'fPN'	0.790	0.988	0.878
'NE'	0.970	0.829	0.894
'aAP'	0.478	0.917	0.629
'VE'	0.948	0.951	0.949
'NP'	1	0.885	0.939
'AE'	0.857	1	0.923
'SP'	0.000	0.167	1
Average	0.923	0.945	0.9258

REFERENCES

[1] S. García and F. Herrera, "Evolutionary undersampling for classification with imbalanced datasets: proposals and taxonomy.," *Evolutionary computation*, vol. 17, no. 3, pp. 275-306, Jan. 2009.

[2] B. X. Wang and N. Japkowicz, "Boosting support vector machines for imbalanced data sets," *Knowledge and Information Systems*, vol. 25, no. 1, pp. 1-20, 2009.

[3] N. Japkowicz, "Learning from Imbalanced Data Sets : A Comparison of Various Strategies," in *Proc. Am. Assoc. for Artificial Intelligence (AAAI) Workshop*, 2000, vol. 68.

[4] C. Vivaracho-pascual and A. Simon-hurtado, "Improving ANN performance for imbalanced data sets by means of the NTIL technique," *IEEE International Join Conference on Neural Networks (IJCNN)*, 2010.

[5] E. M. Imah, W. Jatmiko, and T. Basaruddin, "Adaptive multilayer generalized learning vector quantization (amglvq) as a new algorithm with integrating feature extraction and classification for arrhythmia heartbeats classification," in *IEEE international Conference on System Man and Cybernetics, Seoul 2012*, 2012.

[6] A. Sato and K. Yamada, "Generalized Learning Vector Quantization," in *Advances in Neural Information Processing Systems 8 Proceedings of the 1995 Conference*, 1996, vol. 7, pp. 423-429.

[7] T. Kohonen, "Learning-Vector Quantization and the Self-Organizing Map," in *Theory and Applications of Neural Networks*, 1992, pp. 235-242.

[8] F.-michael Schleif, T. Villmann, B. Hammer, P. Schneider, and M. Biehl, "Generalized derivative based kernelized Learning Vector Quantization," *Springer, Intelligent Data Engineering and Automated Learning – IDEAL 2010 Lecture Notes in Computer Science*, vol. 6283, pp. pp 21-28, 2010.

[9] K. Marika, B. Hammer, M. Biehl, and T. Villmann, "Neurocomputing Functional relevance learning in generalized learning vector quantization," vol. 90, pp. 85-95, 2012.

[10] E. Matul I., I. M. A. Setiawan, a. Febrian, and W. Jatmiko, "Arrhythmia heart beats classification using mahalanobis Generalized Learning Vector Quantization (Mahalanobis GLVQ)," *2011 International Symposium on Micro-NanoMechatronics and Human Science*, pp. 355-360, Nov. 2011.

[11] E. M. Imah, F. Al Afif, M. Ivan Fanany, W. Jatmiko, and T. Basaruddin, "A comparative study on Daubechies Wavelet Transformation, Kernel PCA and PCA as feature extractors for arrhythmia detection using SVM," in *TENCON 2011-2011 IEEE Region 10 Conference*, 2011, pp. 5–9.

[12] W. Jatmiko, W. P. Nulad, E. M. I, and I. M. A. Setiawan, "Heart Beat Classification Using Wavelet Feature Based on Neural Network," vol. 10, no. 1, pp. 17-26, 2011.

[13] F. Sufi and I. Khalil, "Diagnosis of cardiovascular abnormalities from compressed ECG: a data mining-based approach.," *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 15, no. 1, pp. 33-9, Jan. 2011.

[14] P. Tadejko and W. Rakowski, "Mathematical morphology based ECG feature extraction for the purpose of heartbeat classification Faculty of

Computer Science," *6th International Conference on Computer Information System and Industrial Management Application (CISIM'07)*, 2007.

[15] E. Sadat, H. Rooteh, Y. Zhang, and Z. Tian, "Comparison of Parallel and Single Neural Networks in Heart Arrhythmia Detection by Using ECG Signal Analysis," in *PHM*, 2011, pp. 1-9.

[16] G. de L. D. Francois, J. Delbeke and and M. Verleysen, "Weighted SVMs and Feature Relevance Assessment in Supervised Heart Beat Classification," *Commun. Comput. Inf. Sci.*, vol. vol. 127, pp. 212–225, 2011.

[17] G. de Lannoy, D. Francois, J. Delbeke, and M. Verleysen, "Weighted conditional random fields for supervised interpatient heartbeat classification," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 1, pp. 241-7, Jan. 2012.

[18] A. L. Goldberger et al., "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, p. 215, 2000.

[19] Luz, E., et all.," ECG-based heartbeat classification for arrhythmia detection: A survey", Computer Methods and Programs in Biomedicine, Vol 127, Pages: 144–164, 2016.

[20] J. Weng, Cheng G, Poon, "A New Evaluation Measure for Imbalanced Datasets," in *Seventh Austraasian Data Mining Conference (AusDM 2008)*, 2008.