A Multi-Stage Ensemble Neural Network for Accurate Differentiation Between False and Real Alarms in Quality Control Charts

Nader S. Santarisi[1]   Mahmoud Barghash[2]*   Ahmad Barghash[3]

[1]*Mechanical and Industrial Engineering Department, Applied Science Private University, Amman 11931, Jordan*
[2]*Industrial Engineering Department, The School of Engineering, The university of Jordan Amman 11942, Jordan*
[3]*Computer Science Department, German Jordan University,*
School of Electrical Engineering and Information Technology, Amman 11180, Jordan
* Corresponding author's Email: mabargha@ju.edu.jo

Abstract: Faults in machines are common and may cause detrimental damage to the products. Also, faulty raw materials may cause less than the required quality. These machines may be automated, and the fault may go undiscovered for somewhile during which many substandard products may have been produced. Thus, it is important to develop techniques to discover the very minute change in the manufacturing process. For this purpose, many sensors and data acquisition systems are added to some of these machines to monitor the working health of the machine. Faulty machines may manifest themselves as shifts in sensors readings that collect data about the manufacturing process status. Algorithms are then developed to analyse the sensors reading and detect this awkward behaviour as early as possible prior to producing faulty products. This is especially important for machines producing at high speed where faults may result in a significant number of unacceptable products. However, algorithms developed to detect this abnormal behaviour may make wrong decisions and classify normal process as abnormal causing false alarms. These false alarms may be more expensive than faulty products as they reduce the productivity due to unnecessary stoppage. This research work developed an algorithm which is an artificial neural network ensemble formed of two artificial neural networks (ANNs) and two Radial basis functions (RBF) networks in addition to a leader artificial neural network. Individual networks may produce high sensitivity to any changes in the process, but at the same time may cause huge amount of stoppage time due to false alarms. The ensemble suggested addresses that problem and makes a final decision as a second stage from the initial stage decisions of several diverse and tunned networks. The individual networks are trained using various population percentages to intentionally cause decision diversity. The networks' performance is tuned by varying the cut-off parameter. We noticed that the performance of the ANN and RBF in detecting abnormal behaviour is sensitive to the population percentages during the training and the cut-off or threshold parameter. Moreover, through evaluating contradictory metrics, mainly sensitivity, specificity, and accuracy, it was found that the suggested ensemble can be tuned to give multiple acceptable solutions. Also, when benchmarking to literature, it was noticed that the average accuracy of the devised ensemble is 99.13 which is comparable to the best methodologies found in the literature.

Keywords: Artificial neural networks, Artificial neural network ensemble, Artificial neural networks tuning, ANN training population percentages, Control chart pattern recognition, False alarms, Radial basis functions.

1. Introduction, background, and research aims

In smart manufacturing systems machines are equipped with many sensors and communication systems that collect data about the manufacturing process status. This is of utmost importance since some machines produce at high speed and the presence of faulty condition -without early detection- may result in a tangible number of substandard or even defective products. Faulty machine conditions may manifest itself in the form of shift patterns in sensors readings. These patterns can be revealed or analysed using statistical Process Control. Statistical

Process Control (SPC) can be employed to reveal several hidden (perhaps abnormal) patterns through drawing and analysis of control charts. As a result, control chart pattern recognition (CCPR) problem has become extremely important problem which has been addressed by a rising number of research work and it included the development of new algorithms for the detection of abnormal patterns or shifts in the sensors readings. The CCPR problem is highly challenging and requires much faster and more intelligent algorithms [1].

Real time detection of non-standard products or performance is of utmost importance. Statistical Process control is useful in that respect to test for persistent quality performance. Quality control charts are graphical SPC tools that are traditionally used for the detection of patterns in the charts [2]. The system designed to detect these patterns must handle vague situations or patterns that can have similarity between "normal" or random and "abnormal" or causative conditions. In such a vague case a decision might be made to stop the machine thinking that it is in the abnormal condition while it is normal. This case is a false alarm. False alarms reduce the productivity of machines. Another problem with vague situations is to decide to continue with the production while the machine has a faulty behaviour which must be corrected. This situation can lead to faulty products. CCPR algorithms can substantially minimize the time and cost to detect an out-of-control process with high accuracy [3]. The most commonly analysed patterns in CCPR are normal (Nor), up-shift (US), down-shift (DS), up-trend (UT), down-trend (DT), cyclic (CYC), and systematic (SYS) patterns. The first mention for normal and abnormal patterns was in 1958 by Western Electric Company [4].

This important problem had been approached using several methodologies that included basic statistical analysis perspective or Bayesian perspective [5-7, 9, 10, 48, 52], soft computing perspective [8], clustering perspective [11-14], regression perspective [8, 15], Support vector machine perspective [8, 16-23, 41, 43], principle component perspective [19-21, 24], decision tree perspective [25-27, 75] Artificial Neural Network perspective (ANN) [2, 8-11, 16, 28-33, 44, 46, 47], convolution neural network (CNN) perspective [34, 49], expert system perspective [5, 35, 36], Radial Basis Function (RBF) neural network [40, 42], and neuro-fuzzy system perspective [14, 37-39]. Some researchers used wavelet pre-analysis prior to chosen artificial intelligence methodology [21, 29, 33], while others used some kind of feature extraction [50] or mix ratio data simulation [74]. A more thorough literature survey for control-charts pattern
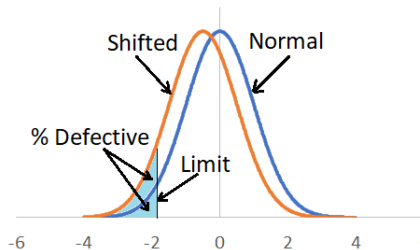

Figure. 1 Change in defective products or material produced with a shift in the process

recognition is available in [45, 73].

Control charts were originally developed by Shewhart in 1920 to address the abnormal process behaviour [51]. It played a critical role in quality engineering. CCPR was specifically developed partly to detect patterns in control charts and partly to categorize these patterns in a way to help in reaching the root cause for abnormal patterns. Huge change in the process, can easily be detected since it will cause a large change in the process sensors readings or product quality measurement. Most of the methodologies developed can easily detect such a major change even the most early developed heuristics can detect these high changes. The real problem, however, is in detecting minor process changes as early as possible. It is also important to differentiate these minor changes from random process variations. Processes include normal variability which has no specific cause, and it cannot be remedied. While minor process changes have causes and can be remedied. The cost incurred in the abnormal process changes may be tangible, in that more material is wasted, more energy is used, more scrap is produced, etc. Even the slightest tangible change in process parameters may incur added costs.

Fig. 1 shows what happens to the % defective when the process encounters a shift. The % defective which represents the products below a specified limit and is calculated as the area under the normal curve. The percentage of defectives will encounter tangible increase under even minor shift in the process mean as illustrated in Fig. 1.

Most of the methodologies devised for pattern recognition will discover major process changes very early in the process, but minor changes might go undetected for a long process run. The length of the run prior to detection of abnormal behaviour is called the Average Run Length (ARL). The ARL is traditional used to quantify the performance of CCPR algorithms. These CCPR algorithms may include tuning parameters that increase the sensitivity to discovers minor changes earlier, however, this increased sensitivity will result in more false alarms because process random behaviour is mixed up with

causative behaviour. Now most of the methodologies did not address this sensitivity mix-up seriously and limited itself to medium to major changes in the process. This work on the other hand builds upon previous work by the authors of this research work [2] which is related to tuning the population composition to improve the performance of the ANN. The population composition effect on the detection accuracy or correct classification was discussed early in the literature as a resampling for imbalanced data [53-68]. The concept of imbalanced population was related to data collected with different numbers of main pattern samples. One pattern may be abundant and the other may be scarce. It was discovered that this variability in the number of pattern samples affected the error rate. It was also discovered that resampling (reducing or increasing the number of pattern sample through copying or reinsertion) can reduce the error rate in the algorithm trained to detect these patterns.

The other approach to handle population composition was through intentionally misbalancing data [2]. The work was related to testing the effect of the training parameters including patterns percentages in control chart on the error rate. In this perspective the intentional imbalance can lead to variation in performance for the ANN. This variation was looked at in two manners: 1) Varying parameters to obtain the best performance, and 2) Varying parameters to obtain neural networks that has different accuracies. This was later used to form an ensemble of different neural networks to obtain more accuracy [69].

The earlier research in pattern recognition in control charts, utilized single type of modelling or clustering analysis techniques to make the decision whether the process is normal or abnormal, such as neural networks, support vector techniques, statistical techniques, expert systems, etc. Moreover, feature extraction had been used to convert that data into features prior to inputting to the classification algorithm. Many of these used threshold-based scoring. A specific window size of the data, whether normal or not, is input to the model and the output of the model is evaluated against a threshold. If the output is above the threshold, it is considered as abnormal otherwise it is normal. These single-model techniques when faced with high similarity of the patterns especially in the case of small shift will result in notifiable awkward outcomes. Many normal patterns may be classified as abnormal and vice versa. Or in the case of evaluating the ARL through window shifting with time, instabilities in the decision will occur. That is, the single model will change its decision with time. Also, if we apply the same pattern

to different classification models, there might be different classification decisions across the different models.

The above within-class decision variability or instability in the decision with time, or even variability across models for the same patterns, opens full potential research work that necessitates group-based techniques to handle these multiple decisions or multiple scores for the same data. This issue can be handled through different techniques including convolutional neural networks, where the input data passes through multiple analysis stages ending with a single vector of class scores [34], or consensus clustering where multiple clustering algorithms are applied to the same data pattern and then consensus matrix is developed [13], or ensemble based techniques where the same data patten is input to multiple neural networks which are trained separately to obtain multiple scores [69]. Following to that step, these multiple scores are used to attain a more reliable and stable decision. This research work is an extension within this current trend in considering the variability in decision across multiple algorithms, and to handle the problem of stabilizing decisions of single models to obtain better classification decisions.

This work is a continuation of the work by using multiple type of neural networks namely Artificial Neural Networks (ANN) and Radial Basis Functions (RBF), in one ensemble to detect shifts in process mean. The ensemble is composed two well-tuned ANN's and two well-tuned RBFs with one leader ANN to make the final decision. We also used in the work 3D graphs to aid the tuning process, and contour maps of the accuracy, sensitivity, and specificity to analyse the behaviour of the ensemble under different settings. In order to achieve this work objectives, Section 2 presents the methodology including ensemble of networks and leader net, patterns generation for training of the ensemble nets, and evaluation metrics. Results are discussed in Section 3 including RBF network tuning, tuning the ensemble leader, and benchmarking the results with literature, while conclusions are found in Section 4.

2. Methodology

2.1 Ensemble of networks and leader net

ANN uses connectivity to transform highly complex nonlinear problems to simple linear problem in the multi-dimensional space. Therefore, ANN's found use in many research areas. The multi-stage ensemble tends to extend that power and simplifies complex problems which may otherwise lead to huge errors [69].

153

For this work we have used several neural networks of various types. Each one is fine tuned to its best performance. The types of networks used are back propagation neural networks and radial basis networks. Fig. 2 shows the methodology for this work. For each type of network, we modify the training population parameters including the population percentages and the cut off point at which the pattern is considered a shift pattern (minimum shift). Training patterns with very small shift may results in high confusion for the network due to closeness of the two patterns. The normal patterns and the shift patterns only differ by the change in the mean. If that change is not significant, there will be a confusion. Thus, during pattern generation for the shift, the mean change should not be less than a certain minimum value. Then, the population percentages are modified to obtain the best performance for each network type. Following that the results for the neural networks are passed to the ensemble for the final training and one decision is made by the leader network which is an ANN network. To stabilize the leader network, we use several decision points from each network.

Fig. 3 shows a schematic for the suggested ensemble. The most noticeable features in the devised ensemble are reducing instability (changing decision with time) through multiple timely decisions per network (3 decisions), using the diversity of the networks to obtain more accurate classification and finally using a threshold-based output network classification for tuning the performance. Thirty consecutive sample points are passed each time to input networks. Then, the ensemble assembles three consecutive decisions per input network to form the input for the leader network. A threshold-based classification is then used; that is, if the output of the leader network is below the threshold, process is considered normal, otherwise a shift in mean is detected. Several neural networks are trained to detect either normal patterns or shift patterns. Population percentages used are 50% normal and 50% shifted, or 10% normal and 90 % shifted. The networks are trained to high accuracy reaching to a minimum sum of squares error. The ANN networks are formed of four layers with the first and the second layer being logsig functions, the third layer being tansig function, and final output layer being purelin function (see Eqs. (1) to (3)). Mathworks MATLAB 2021 is used. The networks are trained using trainlm training algorithm in MATLAB.

$$Logsig(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

$$Purelin(x) = b + x \qquad (2)$$

$$Tansig(x) = sign\left(\frac{2}{1+e^{-2x}} - 1\right) \qquad (3)$$
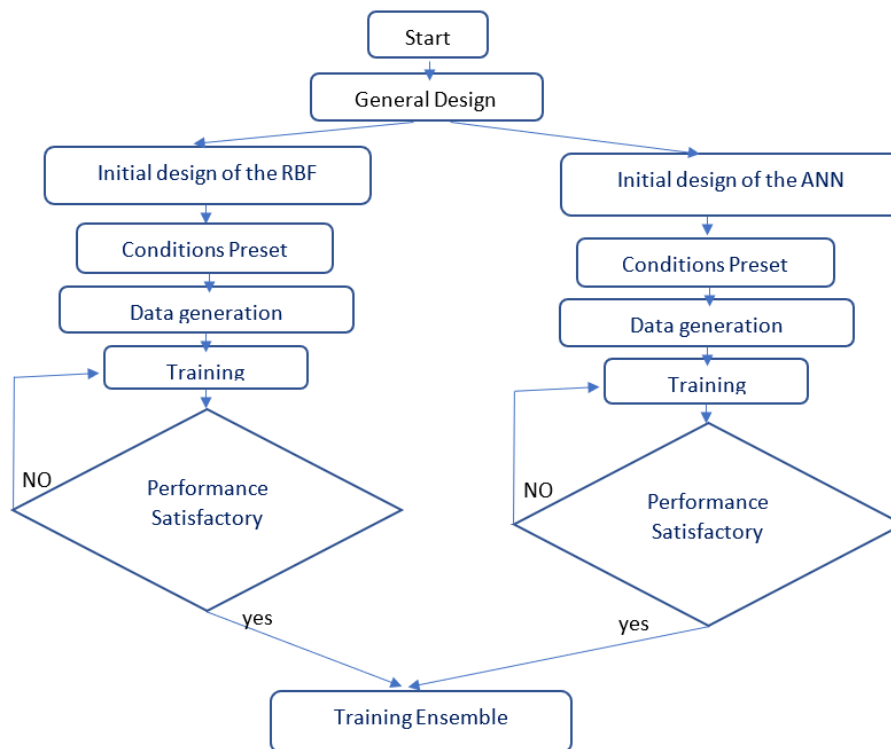

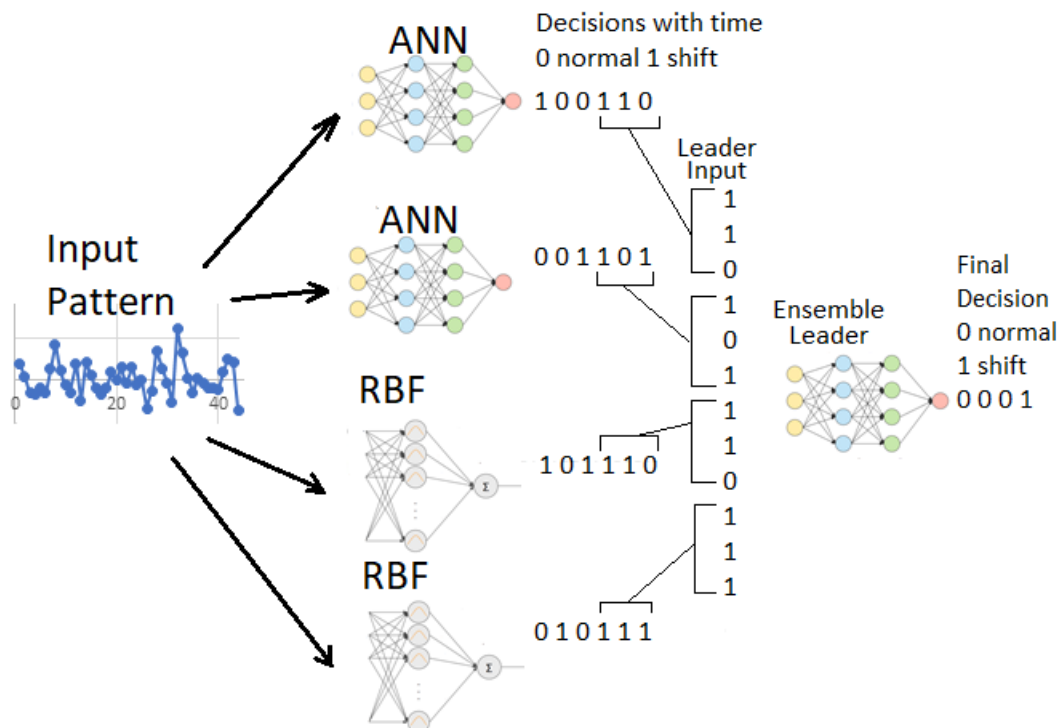Figure. 2 General methodology

Figure. 3 A schematic of the suggested ensemble

The threshold level classification is the most used classification for the output of the neural nets among the researchers. It was tuned in this work to produce better than the benchmark. The benchmark is usually the Shewhart X-bar chart performance levels which has an ARL equal to 370 for the normal process and an ARL of 155 for the 0.5 sigma shift in the mean. Any developed algorithm is considered better if it produces an ARL above 370 for the normal process and an ARL less than 155 for the 0.5 sigma shift in the mean.

2.2 Patterns generation for training of the ensemble nets

The following represents the model for generating the normal and the shifted patterns.

$$y(i) = n(i) + d \qquad (4)$$

Where $y(i)$ is the shifted process signal, $n(i)$ is the normal process with mean μ and standard deviation of σ, and d is the shift size as a multiple of σ. The shift values can be expressed in a standardized form: as - 3,-2.5,-2,-1.5,-1,-0.5, 0.5,1,1.5,2,2.5 and 3. Furthermore, $y(i)$ is standardized into $Y(i)$ using the following equation.

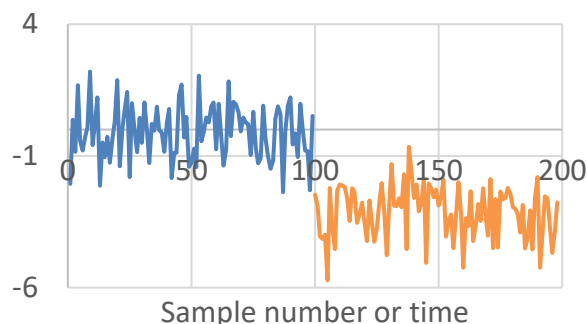$$Y(i) = \frac{y(i) - \mu}{\sigma} \qquad (5)$$

Figure. 4 A sample chart showing a shift in the mean at sample number 100

Fig. 4 shows a sample chart with a mean shift occurring in sample point number 100. Prior to the shift the mean is zero after the shift the mean becomes -3.

2.3 Evaluation metrics

In this work, we will use two types of metrics: detection speed metric, which is the ARL, and classification metrics. The ARL is the number of samples testing negative (normal) while the process is having problems manifested as a shift in the mean. The classification metrics are based upon counting the number of right classification and wrong classification. The following counts are made: True Positive (TP), False Negative (FN) and these two are evaluated for abnormal processes (shift occurring).

Truly Negative (TN) and False Positive (FP) and these are evaluated for a normal process (no shift case). We can then define the following measures [70]:

$$Sensitivity = \frac{TP}{TP+FN} \qquad (6)$$

$$Specificity = \frac{TN}{TN+FP} \qquad (7)$$

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \qquad (8)$$

The sensitivity is the proportion of true shift detected patterns to the total shifted patterns. It shows how good the test is at detecting a shift. Specificity is the proportion of the true normal detected patterns relative to all normal patterns. It suggests how good the test is at identifying normal (negative) condition. Accuracy is the proportion of true results, either true positive or true negative, in a population. It measures the degree of veracity of a diagnostic test on a condition. A test with high sensitivity tents to capture all possible positive conditions without missing anyone. Thus, a test with high sensitivity is often used to screen for shifted patterns or abnormal behavior. The numerical value of specificity represents the probability that a normal process will be classified as normal [70]. These measures are contradictory, in that one algorithm tuned setting will increase the sensitivity but will reduce the specificity. We need to find a tuned setting that will best compromise between these measures.

3. Results

This section discusses the tuning of the networks and the effect of the cut off parameter for various shifts sizes. This is important as an initial stage to the development of the ensemble leader network. The ensemble leader network is then tuned using the cut off parameter to obtain the best possible performance.

3.1 RBF network tuning

Cut off parameter or threshold parameter is extremely important for classification. Usually, the output of the neural network is analogue, and we need to set this parameter such that if the output of the network is above the threshold it is classified as abnormal, otherwise it is considered as normal. Fig. 5 is a 3D plot for the average run length of RBF network established with population percentage 50% normal and 50% shifted. The X-axis is the cut-off parameter. As that parameter is reduced, less patterns are classified as normal, and the average run length is reduced. The y axis in Fig. 5 is the amount of standardized shift. As the shift in the mean increase, the network will easily discover the shift and the average run length is reduced. We need now to compromise between the need to discover shifts early and the false alarms. The false alarm occurs when the shift is zero, and yet the process is classified as abnormal. Tuning the neural here means selecting the right cut off. It is useless to select the cut off for this network below 0.8. Below this value the average run length is low for normal cases when the shift is zero. At a threshold of 0.8 the ARL for shift size zero is nearly 2800. Which is excellent and much better than the benchmark of 370 obtained through traditional techniques. But, for small shift sizes of 0.5 we have an ARL of nearly 2000 which is a very high value, and the system fails to discover this small shift early.

Fig. 6 plots the ARL for the RBF network trained with population percentage 90% normal and 10% shifted. The difference in the two Fig. 5 and 6, stresses the importance of population percentages tuning. This RBF network performance has an ARL of nearly 6000 for no-shift case (shift=zero) while at a shift size of 0.5, it has an ARL of nearly a 1000. Although the performance in discovering small shift is not acceptable, it still has the advantage of very low false alarms rate.

3.2 RBF tuning the ensemble leader

Fig. 7 shows the tuning 3D curve for the leader network of the ensemble. It can be clearly noticed that the ensemble is providing better results than individual networks. When selecting the right cut off, there are three regions, cut off valve below 0.3 where all patterns are classified as abnormal, and the false alarms are excessive. The next region is for cut off having any value between 0.3 and 0.6, and the process in this case has excellent high-speed detection of even the smallest shift, and low number of false alarms with an ARL of 1003. The last region has the lowest false alarms with an ARL of 7016 for the no shift case (shift =0) and slightly high ARL for small shifts and excellent performance for detecting the high shift case.

Table 1 shows the details for Fig. 7 where the ARL sizes for different shift sizes and selected cut-off is evaluated. The recommended two tunning cases are highlighted one for excellent low false alarms and excellent high and moderate shift detection. The other highlighted is for moderate false alarm and excellent detection of even the smallest shift.

Table 1. Tabulated results for the ensemble

Average Run Length		Standardized Shift Size						
		0	0.5	1	1.5	2	2.5	3
	0.1	2	2	2	2	1.9	1.9	1.8
	0.2	48.1	24.1	6.3	5.9	3.5	3.3	2.2
Cut off or threshold	0.3	1003.5	63.1	10.8	9.3	5.9	4.9	2.6
	0.4	1003.5	63.1	10.8	9.3	5.9	4.9	2.6
	0.5	1003.5	63.1	10.8	9.3	5.9	4.9	2.6
	0.6	7016.3	812	239.2	46.4	25.5	9.7	3.6
	0.7	7016.3	812	239.2	46.4	25.5	9.7	3.6
	0.8	7016.3	812	239.2	46.4	25.5	9.7	3.6
	0.9	7016.3	812	239.2	46.4	25.5	9.7	3.6
	1	7016.3	812	239.2	46.4	25.5	9.7	3.6

Fig. 8 shows the "filled contour" for the count classification metrics, namely accuracy, sensitivity, and specificity. The first impact of such a chart indicates contradictory effect of the tuning parameter on these important objectives or measures. This in turn makes the selection of the tuning parameter difficult. Let us elaborate in that as follows. The specificity is the simplest of all charts, as we need to select the cut off parameter above 0.2 to provide acceptable performance for all shift sizes. And this is independent of the shift size. This is logical as we only take into consideration the normal patterns. Logically speaking the sensitivity is best when selecting the tuning parameter as low as possible. This is well shown by the chart. But it is also affected by the shift size. The less the shift size, the more it can get confused with a normal pattern and thus requires lower cut off value to be distinguished better.

This requirement contradicts that of the sensitivity. The accuracy on the hand shows more complex behavior than that of the sensitivity or the specificity. This is due to considering both normal and abnormal patterns. The higher the shift size, the higher the accuracy. We also notice that if we select the cut off parameter extremely close to training limits of 1 or zero the accuracy deteriorates.

Fig. 9 shows a comparison between one of the selected tuned settings and the Shewhart X-bar chart performance levels. The new ensemble appreciably outperforms the benchmark. In that, the number of false alarms will be much less, as the ARL for zero shift size or normal performance is much higher. Also, a 0.5 sigma shift in the mean is detected earlier by the new ensemble.
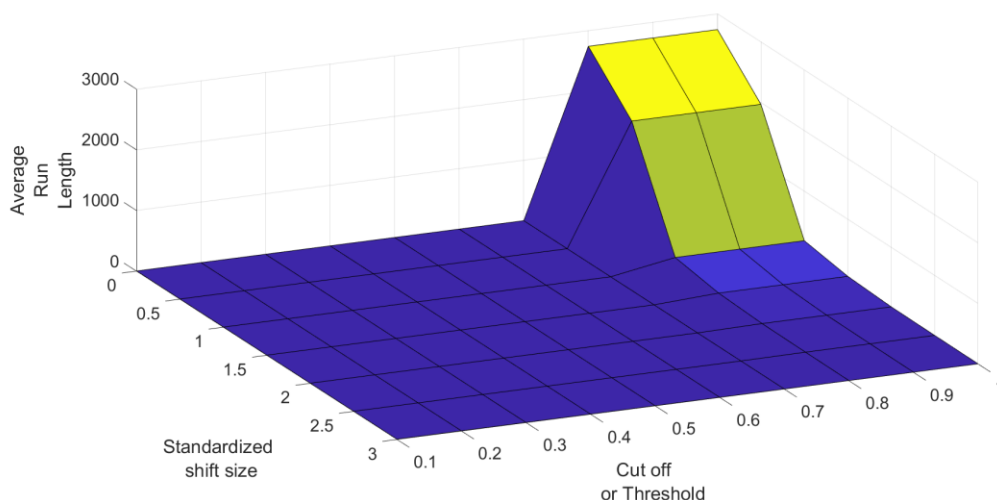

Figure. 5 ARL versus standardizes shift size and cut for an RBF trained for 50% normal 50% shifted population
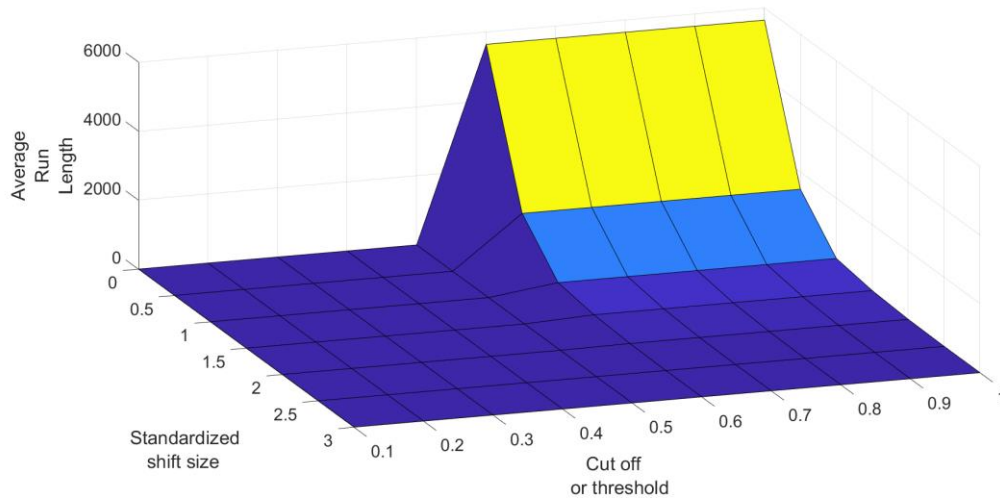
Figure. 6 ARL versus standardizes shift size and cut for an RBF trained for 90% normal 10% shifted population
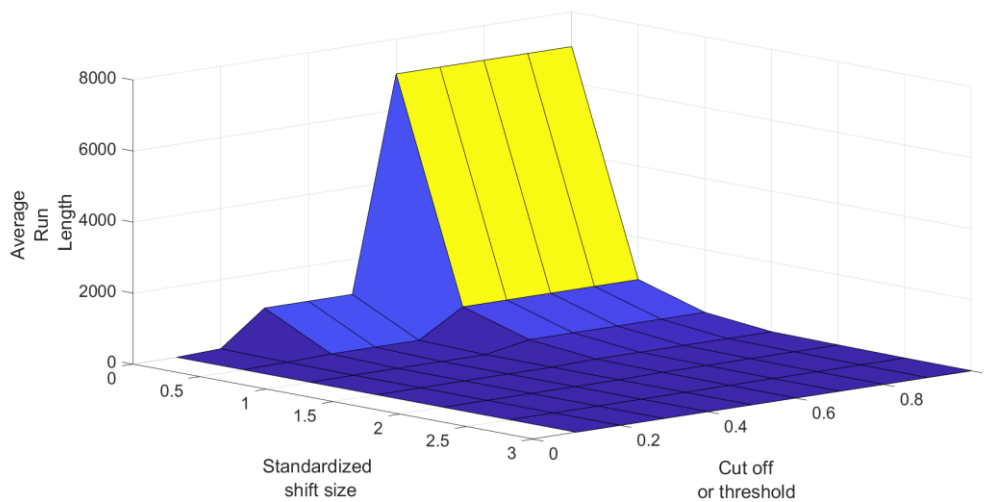

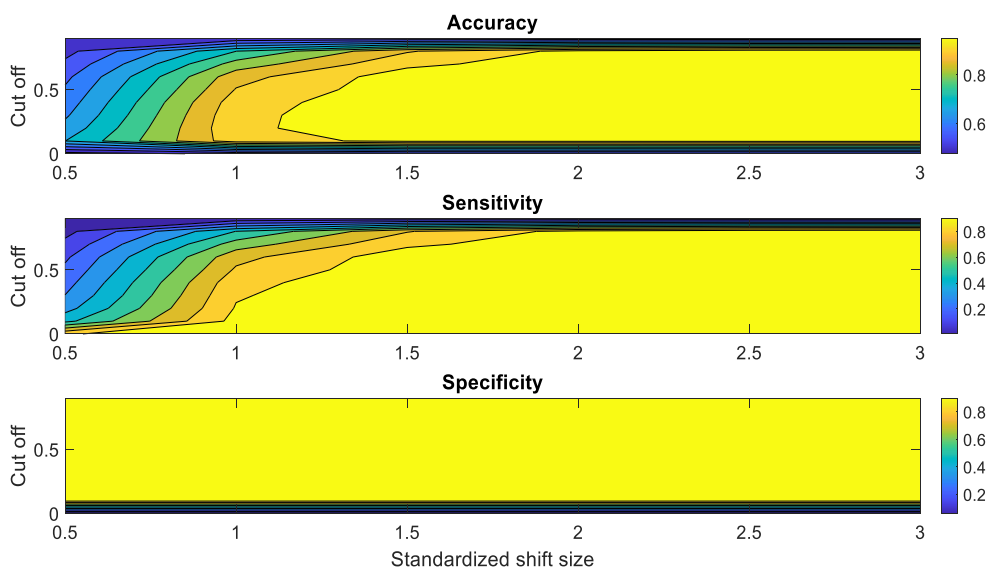Figure. 7 The ensemble leader ANN tuning results in 3D

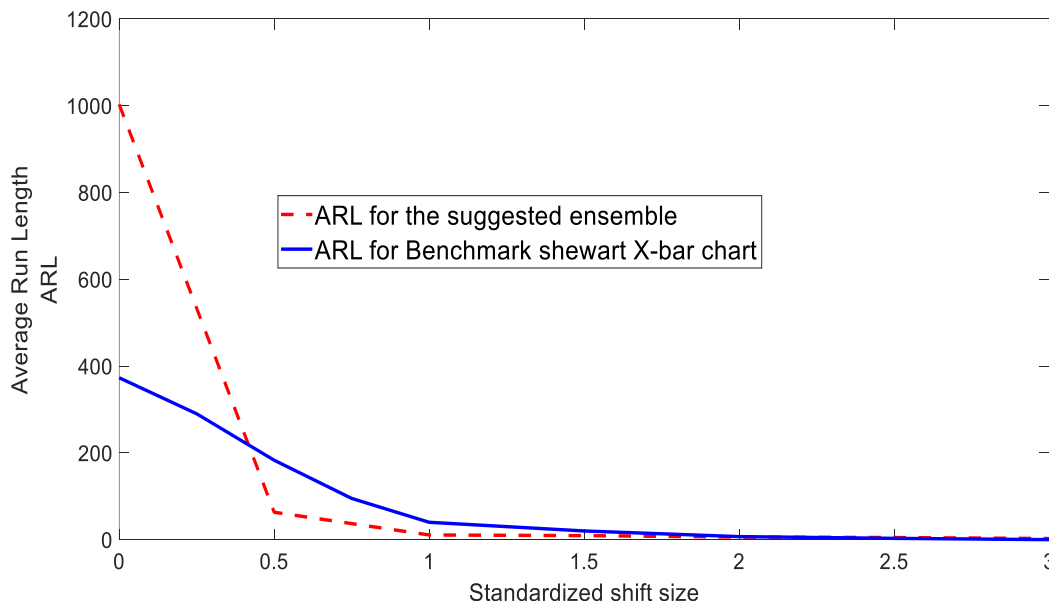Figure. 8 Accuracy, sensitivity, and specificity for the leader network

Figure. 9 A comparison of the ensemble performance to the X-bar chart benchmark

Table 2. Performance comparison of different methods: Modified from [72]

Reference	CRR%
Guh and Tannock (1999)	94.38
Hassan et al. (2003)	96.80
Assaleh and Al-Assaf (2005)	97.22
Cheng and Ma (2008)	95.58
Zan et al. (2010)	95.00
Ranaee and Ebrahimzadeh (2013)	99.15
Kao et al. (2016)	98.94
Zhou et al. (2018)	99.28
Addeh et al. (2018)	99.63
Zan et al. (2020)	99.30
This work	99.13

3.3 Benchmarking the results with literature

We have compared the results of this work with literature using shift sizes of 1.5σ to 3σ, and with shift starting at time zero, same as that used by comparable research. The measure used was the accuracy sometimes referred as average accuracy or correct recognition ratio (CRR) [72]. Table 2 shows the comparison of the results of this work with literature.

Few literature reference were available for low shift sizes so benchmark with literature is difficult in that respect. Also, few literature adopted the ARL measure although it is the most accurate measure in the quality control standard literature.

4. Conclusions

This work developed an ensemble to detect shifts in the process mean which may detect faulty production. The ensemble is formed of two ANN's and two RBF networks. They were trained with population percentages of 10% normal: 90% shifted and 50% normal: 50% shifted. The cut off for each network is tuned to obtain the best performance. The ensemble has a leader network to give the final decision. The following was this work conclusions:

1- The ANN and RBF is sensitive to both the population percentage and the cut off parameters.
2- The ensemble performed better than individual neural networks in terms of the compromise between the false alarms and abnormal behavior.
3- The tunning parameters should be chosen to compromise between the accuracy, sensitivity and specificity and is difficult to optimize all at the same time.
4- The developed ensemble suggested more acceptable settings that will enable users to select what is best for them. This can be interpreted as an added flexibility.
5- The results of this work showed a positive comparison with respect to best-performing models in related literature.

Conflicts of Interest

The authors declare no conflict of interest in publishing this work.

Author Contributions

The authors jointly worked together in formulation, implementation, and programming of the code, in addition to analysis and writing.

References

[1] D. Fuqua and T. Razzaghi, "A cost-sensitive convolution neural network learning for control chart pattern recognition", *Expert Systems with Applications*, Vol. 150, 113275, pp. 1-17, 2020.

[2] M. Barghash and N. S. Santarisi, "Pattern recognition of control charts using artificial neural networks—analyzing the effect of the training parameters", *Journal of Intelligent Manufacturing*, Vol. 15, No. 5, pp. 635-644, 2004.

[3] D. C. Montgomery, "Statistical quality control", *Wiley Global Education*, 2012.

[4] Western Electric Company, "Statistical quality control handbook", *Western Electric*, 1958.

[5] J. A. Swift, "Development of a knowledge based expert system for control chart pattern recognition and analysis", *Ph.D. thesis, Oklahoma State University*, 1987.

[6] J. H. Yang and M. S. Yang, "A control chart pattern recognition system using a statistical correlation coefficient method", *Computers & Industrial Engineering*, Vol. 48, pp. 205-221, 2005.

[7] M. K. Naeini, M. Owlia, and M. Fallahnezhad, "A bayesian approach for recognition of control chart patterns", *International. Journal of Industrial Engineering*, Vol. 23, 2012.

[8] Y. E. Shao and C. C. Chiu, "Applying emerging soft computing approaches to control chart pattern recognition for an spc–epc process", *Neurocomputing*, Vol. 201, pp. 19-28, 2016.

[9] N. A. Rahman, I. Masood, M. N. A. Rahman, and N. F. Nasir, "Control chart pattern recognition in metal-stamping process using statistical features-ann", *Journal of Telecommunication. Electronic and Computer Engineering*, Vol. 9, pp. 5-9, 2017.

[10] C. S. Cheng, K. K. Huang, and P. W. Chen, "Recognition of control chart patterns using a neural network-based pattern recognizer with features extracted from correlation analysis", *Pattern Analysis and Applications*, Vol. 18, No. 1, pp. 75-86, 2015.

[11] A. Ebrahimzadeh, J. Addeh, and Z. Rahmani, "Control chart pattern recognition using k-mica clustering and neural networks", *ISA Transactions*, Vol. 51, pp. 111-119, 2012.

[12] A. Khormali and J. Adde, "A novel approach for recognition of control chart patterns: Type-2 fuzzy clustering optimized support vector machine", *ISA Transactions*, Vol. 63, pp. 256-264, 2016.

[13] S. Haghtalab, P. Xanthopoulos, and K. Madani, "A robust unsupervised consensus control chart pattern recognition framework", *Expert Systems with Applications*, Vol. 42, pp. 6767-6776, 2015.

[14] M. Zaman and A. Hassan, "Improved statistical features-based control chart patterns recognition using anfis with fuzzy clustering", *Neural Computing and Applications*, pp. 1-15, 2019.

[15] B. Mandel, "The regression control chart", *Journal of Quality Technology*, Vol. 1, pp. 1-9, 1969.

[16] C. S. Cheng and H. P. Cheng, "Identifying the source of variance shifts in the multivariate process using neural networks and support vector machines", *Expert Systems with Applications*, Vol. 35, pp. 198-206, 2008.

[17] F. K. Wang, Bizuneh, and X. B. Cheng, "One-sided control chart based on support vector machines with differential evolution algorithm", *Quality and Reliability Engineering International*, Vol. 35, No. 6, pp. 1634-1645, 2019.

[18] C. Zhao, C. Wang, L. Hua, X. Liu, Y. Zhang, and H. Hu, "Recognition of control chart pattern using improved supervised locally linear embedding and support vector machine", *Procedia Engineering*, Vol. 174, pp. 281-288, 2017.

[19] L. J. Kao, T.S. Lee, and C. J. Lu, "A multi-stage control chart pattern recognition scheme based on independent component analysis and support vector machine", *Journal of Intelligent Manufacturing*, Vol. 27, pp. 653-664, 2016.

[20] C. J. Lu, Y. E. Shao, and C. C. Li, "Recognition of concurrent control chart patterns by integrating ica and svm", *Applied Mathematics & Information Sciences*, Vol. 8, p. 681, 2014.

[21] L. Y. Zhong, Z. H. Lie, L. Y. Ju, and J. J. Gang, "Hybrid patterns recognition of control chart based on wa-pca", *International Journal of Control and Automation*, Vol. 7, pp. 91-98, 2014.

[22] P. Xanthopoulos and T. Razzaghi, "A weighted support vector machine method for control chart pattern recognition", *Computers & Industrial Engineering*, Vol. 70, pp. 134-149, 2014.

[23] S. Du, D. Huang, and J. Lv, "Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines", *Computers & Industrial Engineering*, Vol. 66, pp. 683-695, 2013.

[24] F. Aparisi, "Hotelling's t2 control chart with adaptive sample sizes", *International Journal of Production Research*, Vol. 34, pp. 2853-2862, 1996.

[25] Z. Bo, L. Beibei, W. Yuwei, and Z. Shengran, "Recognition of control chart patterns in auto-

correlated process based on random forest", In: *Proc. of 2018 IEEE International Conference on Smart Manufacturing, Industrial & Logistics Engineering*, pp. 53-57, 2018.

[26] C. H. Wang, R. S. Guo, M. H. Chiang, and J. Y. Wong, "Decision tree based control chart pattern recognition", *International Journal of Production Research*, Vol. 46, pp. 4889-4901, 2008.

[27] R. S. Guh, "A hybrid learning-based model for on-line detection and analysis of control chart patterns", *Computers & Industrial Engineering*, Vol. 49, pp. 35-62, 2005.

[28] A. Shaban and M. A. Shalaby, "A double neural network approach for the identification and parameter estimation of control chart patterns", *International Journal of Quality Engineering and Technology*, Vol. 3, pp. 124-138, 2012.

[29] J. S. Kim, C. S. Park, J. G. Baek, and S. S. Kim, "Control chart pattern recognition using wavelet based neural networks", *International Journal of International Science Index*, Vol. 6, pp. 1044-1048, 2012.

[30] M. H. Awadalla, I. Ismaeil, and M. A. Sadek, "Spiking neural network-based control chart pattern recognition", *Journal of Engineering and Technology Research*, Vol. 3, pp. 5-15, 2011.

[31] G. A. Pugh, "Synthetic neural networks for process control", *Computers & Industrial Engineering*, Vol. 17, pp. 24-26, 1989.

[32] C. S. Cheng, "A neural network approach for the analysis of control chart patterns", *International Journal of Production Research*, Vol. 35, pp. 667-697, 1997.

[33] H. P. Cheng and C. S. Cheng, "Control chart pattern recognition using wavelet analysis and neural networks", *Journal of Quality*, Vol. 16, pp. 311-321, 2009.

[34] T. Zan, Z. Liu, H. Wang, M. Wang, and X. Gao, "Control chart pattern recognition using the convolutional neural network", *Journal of Intelligent Manufacturing*, Vol. 31, pp. 703-716, 2020.

[35] S. Alexander, "The application of expert systems to manufacturing process control", *Computers & Industrial Engineering*, Vol. 12, pp. 307-314, 1987.

[36] C. S. Cheng and N. F. Hubele, "Design of a knowledge-based expert system for statistical process control", *Computers & Industrial Engineering*, Vol. 22, pp. 501-517, 1992.

[37] O. E. Farissi and A. Moudden, "Application of neuro-fuzzy in the recognition of control chart patterns", *International Journal of Computer Applications*, Vol. 975, 8887, 2017.

[38] A. Nikpey, S. Mirzaei, M. Pourmandi, and J. Addeh, "Identification of the control chart patterns using the optimized adaptive neuro-fuzzy inference system", *International Journal of Modern Education and Computer Science*, Vol. 6, 16, 2014.

[39] J. Chen and Y. Liang, "Development of fuzzy logic-based statistical process control chart pattern recognition system", *The International Journal of Advanced Manufacturing Technology*, Vol. 86, pp. 1011-1026, 2016.

[40] A. Addeh, A. Khormali, and N. A. Golilarz, "Control chart pattern recognition using rbf neural network with new training algorithm and practical features", *ISA Transactions*, Vol. 79, pp. 202-216, 2018.

[41] R. B. Chinnam, "Support vector machines for recognizing shifts in correlated and other manufacturing processes", *International Journal of Production Research*, Vol. 40, pp. 4449-4466, 2002.

[42] D. F. Cook and C. C. Chiu, "Using radial basis function neural networks to recognize shifts in correlated manufacturing process parameters", *IIE Transactions*, Vol. 30, pp. 227-234, 1998.

[43] M. Davy, F. Desobry, A. Gretton, and C. Doncarli, "An online support vector machine for abnormal events detection", *Signal Processing*, Vol. 86, pp. 2009-2025, 2006.

[44] T. T. E. Midany, M. E. Baz, and M. A. Elwahed, "A proposed framework for control chart pattern recognition in multivariate process using artificial neural networks", *Expert Systems with Applications*, Vol. 37, pp. 1035-1042, 2010.

[45] W. Hachicha and A. Ghorbel, "A survey of control-chart pattern-recognition literature (1991–2010) based on a new conceptual classification scheme", *Computers & Industrial Engineering*, Vol. 63, pp. 204-222, 2012.

[46] H. Hwarng, "Multilayer perceptions for detecting cyclic data on control charts", *International Journal of Production Research*, Vol. 33, No. 11, pp. 3101-3117, 1995.

[47] K. Y. Jang, K. Yang, and C. Kang, "Application of artificial neural network to identify non-random variation patterns on the run chart in automotive assembly process", *International Journal of Production Research*, Vol. 41, No. 6, pp. 1239-1254, 2003.

[48] M. K. Naeini, M. Owlia, and M. Fallahnezhad, "A bayesian approach for recognition of control chart patterns", *International Journal of Industrial Engineering*, Vol. 23, pp. 223-230, 2012.

[49] D. Fuqua and T. Razzaghi, "A cost-sensitive convolution neural network learning for control chart pattern recognition", *Expert Systems with Applications*, Vol. 150, 113275, 2020.

[50] G. D. Pelegrina, L. T. Duarte, and C. Jutten, "Blind source separation and feature extraction in concurrent control charts pattern recognition: Novel analyses and a comparison of different methods", *Computers & Industrial Engineering*, Vol. 92, pp. 105-114, 2016.

[51] W. A. Shewhart, "Economic control of quality of manufactured product", *ASQ Quality Press*, 1931.

[52] J. H. Yang and M. S. Yang, "A control chart pattern recognition system using a statistical correlation coefficient method", *Computers & Industrial Engineering*, Vol. 48, pp. 205-221, 2005.

[53] Y. Sun, A. K. Wong, and M. S. Kamel, "Classification of imbalanced data: A review", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 23, pp. 687-719, 2009.

[54] A. Estabrooks, "A Combination Scheme for Inductive Learning from Imbalanced Data Sets", *MCS Thesis, Faculty of Computer Science, Dalhousie University*, 2000.

[55] L. K. Hansen and P. Salamon, "Neural network ensembles", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 10, pp. 993-1001, 1990.

[56] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive mixtures of local experts", *Neural Computation*, Vol. 3, pp. 79-87, 1991.

[57] N. Japkowicz and S. Stephen, "The class imbalance problem: a systematic study", *Intelligent Data Analysis*, Vol. 6, No. 5, pp. 429-450, 2002.

[58] N. Japkowicz, C. Myers, and M. Gluck, "A novelty detection approach to classification", In: *Proc. of the Fourteenth Joint Conference on Artificial Intelligence*, pp. 518-523, 1995.

[59] M. Kubat and S. Matwin, "Addressing the curse of imbalanced data sets: one-sided sampling", In: *Proc. of the Fourteenth International Conference on Machine Learning*, pp. 179-186, 1997.

[60] M. Kubat, R. Holte, and S. Matwin, "Machine learning for the detection of oil spills in satellite radar images", *Machine Learning*, Vol. 30, pp. 195-215, 1998.

[61] D. Lewis and W. Gale, "Training text classifiers by uncertainty sampling", In: *Proc. of the Seventh Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 1994.

[62] C. Ling and C. Li, "Data mining for direct marketing: problems and solutions", In: *Proc. of KDD-98*, 1998.

[63] M. Pazzani, C. Merz, P. Murph, K. Ali, T. Hume, and C. Brunk, "Reducing misclassification costs", In: *Proc. of the Eleventh International Conference on Machine Learning*, pp. 217-225, 1994.

[64] F. Provost and T. Fawcett, "Robust classification for imprecise environments", *Machine Learning Journal*, Vol. 42, No. 3, 2001.

[65] S. Scott and S. Matwin, "Feature engineering for text classification", In: *Proc. of the Sixteenth International Conference on Machine Learning*, pp. 379-388, 1999.

[66] Y. Shimshoni and N. Intrator, "Classifying seismic signals by integrating ensembles of neural networks", *IEEE Transactions on Signal Processing, Special Issue on Neural Networks*, Vol. 46, No. 5, 1998.

[67] S. Weiss and I. Kapouleas, "An empirical comparison of pattern recognition, neural nets and machine learning methods", In: *Readings in Machine Learning. Edited by J. W Shavlik and T. G. Dietterich*, Morgan-Kauffman, San Maeto, CA., 1990.

[68] G. Weiss and F. Provost, "Learning when training data are costly: the effect of class distribution on tree induction", *Journal of Artificial Intelligence Research*, Vol. 19, pp. 315-354, 2003.

[69] M. Barghash, "An effective and novel neural network ensemble for shift pattern detection in control charts", *Computational Intelligence and Neuroscience*, pp. 1-9, 2015.

[70] W. Zhu, N. Zeng, and N. Wang, "Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations", *NESUG proceedings: health care and life sciences*, 19, 67, 2010

[71] Z. Cheng and Y. Z. Ma, "A research about pattern recognition of control chart using probability neural network", In: *Proc. of ISECS International Colloquium on Computing, Communication, Control, & Management*, 2008.

[72] T. Zan, Z. Liu, H. Wang, M. Wang, and X. Gao, "Control chart pattern recognition using the convolutional neural network", *Journal of Intelligent Manufacturing*, Vol. 31, No. 3, pp. 703-716, 2020.

[73] P. H. Tran, A. A. Nadi, T. H. Nguyen, K. D. Tran, and K. P. Tran, "Application of Machine Learning in Statistical Process Control Charts: A Survey and Perspective", In: *Proc. of Control*

Charts and Machine Learning for Anomaly Detection in Manufacturing, pp. 7-42, 2022.

[74] R. Ünlü, "A robust data simulation technique to improve early detection performance of a classifier in control chart pattern recognition systems", *Information Sciences*, Vol. 548, pp. 18-36, 2021.

[75] M. Zaman and A. Hassan, "Fuzzy Heuristics and Decision Tree for Classification of Statistical Feature-Based Control Chart Patterns", *Symmetry*, Vol. 13, No. 1, p. 110, 2021.