RESEARCH PROGRESS IN MECHANICAL AND MANUFACTURING ENGINEERING
e-ISSN: 2773-4765

RPMME

Vol. 5 No. 1 (2024) 358-365
https://publisher.uthm.edu.my/periodicals/index.php/rpmme

Setting Parameters of Control Chart Pattern Classifier for Effective Patterns Recognition

Muhammad Ammar Abdul Halim[1], Ibrahim Masood[1*]

[1] *Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, 86400, Parit Raja, Johor, MALAYSIA*

*Corresponding Author: ibrahim@uthm.edu.my
DOI: https://doi.org/10.30880/rpmme.2024.05.01.041

Article Info

Received: 15 January 2024
Accepted: 15 July 2024
Available online: 15 Sept. 2024

Keywords

Artificial neural network, control chart pattern recognition, design of experiments, abnormal patterns, statistical process control

Abstract

This research study examines the dynamic field of control chart pattern recognition, which has seen tremendous growth in research over the last few decades. The field has seen advancements in efforts to improve the accuracy of artificial neural network (ANN) based classifier for control chart patterns recognition, including statistical features, wavelet-based denoising techniques for input representation, and the development of both integrated and modular recogniser designs. This study applies the design of experiments (DOE) methodology to select the setting parameters for the ANN-based classifier. This approach is meant to aid in improving and optimising pattern recognition systems within the context of statistical process control. This study's main goal is to assess an ANN classifier's effectiveness in identifying different kinds of abnormal patterns within statistical process control. MATLAB programming software generated the simulated control chart pattern (CCP) samples. The multilayer perceptron neural network model was utilized as the CCP classifier. The setting parameters were analysed and optimized using the design of the experiment technique. The result from this study shows that a low window size is enough to achieve 100% for normal patterns and cyclic patterns. From the New Set 3 that has parameter low recognition window size (20), high amount of training dataset for the normal pattern (1500), low amount of training dataset for shifts pattern (50), high-quality data abnormal pattern for shifts patterns (0.1) and low-quality data abnormal pattern for trends pattern (0.002), it can conclude that the factor is superior for recognising shifts patterns also known as balance recognition.

1. Introduction

This study focuses on improving the performance and accuracy of control chart pattern recognition. Control charts are widely used in statistical process control (SPC) to monitor and assess the stability of processes over time. They help identify deviations from the norm, detect potential issues, and maintain process quality. In the context of control chart pattern recognition, it is essential to identify and classify various patterns that can occur in control charts, such as normal, shifts, trends, and cycles. These patterns can be indicative of underlying process changes, machine malfunctions, or other factors affecting the stability of the monitored process.

The project will involve extensive data analysis, algorithm development, and performance evaluation. Historical control chart data from various industries will be collected and used to train the pattern classifier. Additionally, a robust validation process will be employed to ensure the reliability and generalization of the developed classifier. By setting the parameters of the control chart pattern classifier, the project aims to enhance the overall effectiveness of pattern recognition, leading to better decision-making processes in industrial settings. The successful implementation of this project can significantly contribute to improved process monitoring and quality control, ultimately resulting in enhanced productivity and reduced waste in manufacturing and other critical domains.

Process variation (PV) monitoring and diagnosis are essential for continual quality improvement in manufacturing industries. Using Statistical Process Control (SPC) tools, the control chart is useful in monitoring PV by identifying the relationship between control chart patterns (CCP) and the sources of abnormal PV. In common practice, interpreting CCP requires an experienced industrial practitioner (such as an engineer, foreman, or technician) with a deep knowledge of manufacturing process behaviour. However, this requirement may not be obtained when a company only has new staff. Improper identification of CCP must be avoided since it can lead to wrong process diagnosis, erroneous decision-making, and increased cost of waste. Waste can be defined in terms of waste of materials, time-consuming finding the root cause of error, unnecessary diagnosis activities, and others.

The study aims to develop a basic control chart pattern recognition framework for identifying abnormal control chart patterns and to select a proper setting parameter for the pattern classifier towards improving the classification accuracy.

1.1 Traditional Control Chart Scheme

Traditional control charting schemes are statistical tools used in quality control to monitor and maintain process stability. These schemes help identify any variations or abnormalities in the process and enable corrective actions to be taken. Here are some commonly used traditional control charting schemes:

a) Shewhart Control Charts: Developed by Walter A. Shewhart, these include X-bar and R charts (monitoring central tendency and variability) and X-bar and S charts (similar to the former but monitoring variability through standard deviation) [1].
b) Cumulative Sum (CUSUM) Charts: Detect small incremental shifts in process mean by plotting cumulative sums of deviations [2].
c) Exponentially Weighted Moving Average (EWMA) Charts: Monitor processes with autocorrelated data or for faster detection of small shifts by assigning more weight to recent observations [3].
d) Moving Average (MA) and Moving Range (MR) Charts: These are used when data is collected in groups or batches, with MA tracking the group average and MR tracking the range within a group.
e) Attribute Control Charts: Monitor discrete data's nonconforming items or defects through charts like p-p-charts (proportion nonconforming), c-charts (number of defects), and np-charts (number of nonconformities in a sample).

These traditional control charting schemes provide a systematic approach to monitoring and maintaining process control, allowing organizations to identify and address variations and improve quality.

1.1.1 Shewhart Control Charts

The Shewhart control chart, created by Dr. Walter A. Shewhart in the 1920s, is a visual aid for overseeing and managing processes. It assesses if a process is steady and within statistical boundaries or if exceptional variations need attention. This tool tracks process data over time and contrasts it with predetermined control limits. Regular use of control charts enables organizations to spot and rectify problems before they cause defects or depart from desired standards, offering a visual depiction of performance and supporting ongoing improvement endeavours.

1.1.2 Control Chart Pattern

Control charts are statistical tools used to monitor and analyse process data over time. They help identify patterns, trends, and variations in the data, allowing for the detection of potential problems or improvements in the process [4]. While control charts can vary in design and application, there are several common patterns that can appear on a control chart. Here are some key control chart patterns:

a) Trend: Data points consistently increasing or decreasing over time indicate a trend pattern. This signals a shift in the mean and warrants further investigation.
b) Cycles: Repetitive patterns suggest periodic variations caused by factors like seasons or time of day. Understanding these cycles aids in predicting and managing process fluctuations.
c) Shift: A sudden change in the process mean signifies a shift pattern, which can be upward or downward. This hints at changes in the process or external factors influencing outcomes [5].
d) Stratification: The division of data into subgroups based on specific characteristics helps identify variations within each subgroup. Comparing these subgroups can reveal differences that are not easily apparent when analyzing the data as a whole [6].

It's important to note that these patterns are not exhaustive, and other patterns specific to the process being analyzed may emerge. Control charts visually represent data and help identify patterns that may require further investigation or action to maintain process stability and quality control.

1.2 Nelson's Run Rule

Nelson run rules are eight rules used to identify patterns in control charts that may indicate process instability or out-of-control conditions. The rules were developed by Lloyd S. Nelson in 1984 and are an extension of the Western Electric rules. The Nelson run rules are useful for identifying potential problems with a process. A few researchers employed just one of Nelson's run rules to boost the effectiveness of the Shewhart Control chart [1]. However, it is important to note that they are not perfect and can sometimes trigger false alarms. It is always important to investigate any patterns identified using the Nelson run rules to determine whether they actually indicate a problem.

1.3 Control Chart Pattern Recognition

Control chart pattern recognition is a technique used in statistical process control (SPC) to identify specific patterns or signals in control charts. Control charts are graphical tools that monitor and analyse process data over time to determine if a process is in control or out of control [7].

The primary purpose of control charts is to distinguish between common cause variation (random variation inherent in a process) and special cause variation (non-random variation caused by specific factors). By identifying patterns or signals in control charts, practitioners can detect the presence of special causes and take appropriate action to investigate and eliminate them [8].

It's important to note that control chart pattern recognition is not a definitive method for diagnosing specific causes of process variation. It serves as an initial signal to investigate and identify potential issues that require further analysis and corrective actions. Domain expertise and a deep understanding of the process under observation are crucial for accurate interpretation and effective problem-solving [9].

1.4 Limitation of Existing Pattern Recognition Schemes

Existing pattern recognition schemes have made significant advancements in various domains, but they also have certain limitations. Some of the common limitations are limited generalization which is pattern recognition algorithms often struggle to generalize well beyond the specific patterns they were trained on. They may perform well on training data but fail to accurately recognize patterns in unseen or slightly different contexts. Overfitting is a constraint that causes the model to become overly particular to the training set and underperform on fresh data [10].

Next, the sensitivity to input variations also relate because many pattern recognition schemes are highly sensitive to variations in input data. Even minor changes in the input pattern, such as noise, occlusion, or changes in scale or orientation, can significantly affect the recognition accuracy. This limitation makes the algorithms less robust and less reliable in real-world scenarios where input data may vary.

Training data requirements also includes because the pattern recognition algorithms typically require a large amount of labelled training data to achieve good performance. Obtaining and annotating such datasets can be time-consuming, expensive, and sometimes infeasible for certain domains or rare patterns. Moreover, biases present in the training data can lead to biased or unfair recognition outcomes. Addressing these limitations is an active area of research, and ongoing efforts aim to develop more robust, interpretable, context-aware, and efficient pattern recognition schemes.

2. Methodology

2.1 Research Question and Solution

<table>
<tr><td colspan="2">Research objective (i): To develop a basic control chart pattern recognition framework for identifying abnormal control chart patterns.</td></tr>
<tr><td>Research question</td><td>Research solution</td></tr>
<tr><td>1. How to model the artificial/simulated CCP?</td><td>- Apply the dynamic process data schemes in modelling SPC samples and patterns.
- Use a programming software to represent a visual representation.</td></tr>
<tr><td colspan="2">Research objective (ii): To select a proper setting parameter for the pattern classifier towards improving the classification accuracy.</td></tr>
<tr><td>2. How to design pattern classifier?</td><td>- Investigate the multilayer perceptron neural network model.</td></tr>
</table>

3. How to find the setting parameter?	- Input setting parameter. - Using design of experiment in MINITAB.

2.2 Modelling Statistical Process Control Samples and Patterns

Several common pattern recognition schemes can be used to identify potential issues when analyzing a baseline control chart. Baseline control charts are used in statistical process control (SPC) to monitor and identify any significant deviations or patterns in a process. They aid in detecting whether a process is operating within acceptable limits or if any special causes of variation are present. Figure 3.1 shows the basic line of CCP recognition schemes.

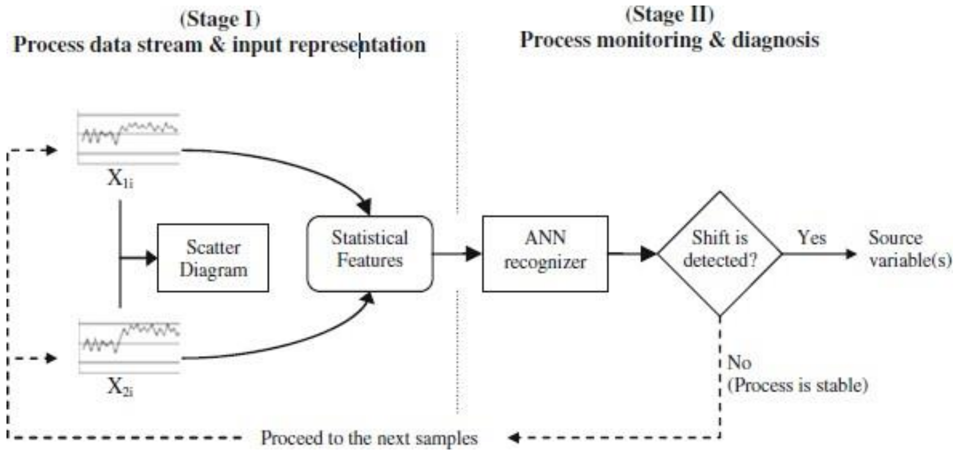

Fig. 1 *Basic CCP recognition schemes [2]*

2.3 Control Chart Patterns

There are six possible categories of SPC as follows are considered in representing the SPC variation in mean shifts:

i.	Normal (N)	: SPC samples remain in-control
ii.	Up-Shift (US)	: SPC samples in upward shifts direction
iii.	Down-Shift (DS)	: SPC samples in downward shifts direction
iv.	Up-Trend (UT)	: SPC samples in upward trends direction
v.	Down-Trend (DT)	: SPC samples in downward trends direction
vi.	Cyclic (CYC)	: SPC samples in cyclic condition

2.4 Recognizer Design

The multilayer-perceptron (MLP) model served as the foundation for the artificial neural network (ANN)-based recognizer, which was then trained using the back-propagation (BPN) technique. Fig. 2 shows ANN recognizer and Fig. 3 from Guh (2007) illustrates a four-layered MLP model.

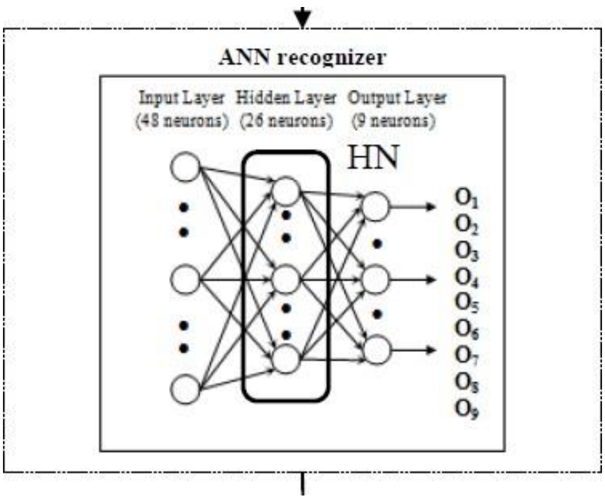

Fig. 2 *Example of ANN recognizer [11]*

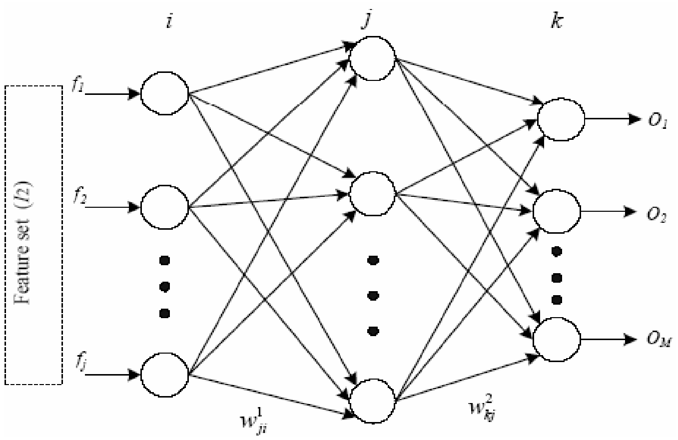

Fig. 3 *Example of MLP neural network structure [12]*

2.5 Setting Parameter

In ANN, setting parameters refer to the choices and configurations made during the construction and training of the network. These parameters significantly impact the network's performance and ability to learn and generalize from data. Four setting parameters were used in these studies: window size, hidden neuron, training dataset, and quality data. Table 1 shows the relationship of the parameter to ANN and an example of the limit level used.

Table 1 *Setting parameter and relationship to ANN [11]*

Design Parameters	**Relationship to ANN**
Recognition Window size (WS)	Size for input layer
Number of Hidden Neuron (HD)	Size of hidden layer
Amount of Training Dataset (TD)	Recognition capability
Quality Data of Abnormal Pattern (QD)	Recognition capability

It's important to note that the selection and tuning of these design parameters are problem-specific and require experimentation and iterative refinement. The optimal choices for window size, hidden neurons, training dataset, and the quality of abnormal patterns depend on the data's characteristics, the problem's complexity, and the neural network's specific goals.

3. Result and Discussion

3.1 Data Analysis

The mean square error (MSE), is a commonly used metric to assess how closely a goal value aligns with the measured error in a simulation. The idea behind the MSE is that a lower MSE indicates a significantly reduced discrepancy between the goal value and the measured error, which lowers the overall percentage of error. More specifically, a higher-performing simulation will have a lower MSE, indicating a closer alignment between the desired outcome and the actual measured error. However, the optimal data will be chosen and summed up later on, though, based on how well the percentage of normal patterns compares to the average of other abnormal patterns.

3.2 Simulation Result

Using the design of the experiment (DOE), the simulation was conducted for 16 variables (partial fractional Type 5) in combination with the selected input data. The input selection parameter is shown in Table 2.

Table 2 *Setting parameters that use and relationship to ANN*

Design Parameters	Relationship to ANN	Low Level	High Level
Recognition Window size (WS)	Size for input layer	20	24
Amount of Training Dataset for Normal (TD_N)	Recognition capability	500	1500
Amount of Training Dataset for Shift (TD_S)		50	100
Amount of Training Dataset for Cyclic (TD_CYC)		50	
Quality Data of Abnormal Pattern for Shift (QD_S)	Recognition capability	0.2	0.1
Quality Data of Abnormal Pattern for Trend (QD_T)		0.002	0.001
Number of Hidden Neurons for ANN Training (HD)	Size of hidden layer	30	

The variable used in the testing was created using the experiment design in MINITAB (Fig. 4). The result of the experiment design is shown in the figure. Through the design of experiments, the input selection parameter is showed -1(low data) and 1(high data).

	StdOrder	RunOrder	CenterPt	Blocks	WS	TD_N	TD_S	QD_S	QD_T
1	8	1	1	1	1	1	1	-1	-1
2	4	2	1	1	1	1	-1	-1	1
3	7	3	1	1	-1	1	1	-1	1
4	2	4	1	1	1	-1	-1	-1	-1
5	13	5	1	1	-1	-1	1	1	1
6	10	6	1	1	1	-1	-1	1	1
7	5	7	1	1	-1	-1	1	-1	-1
8	15	8	1	1	-1	1	1	1	-1
9	6	9	1	1	1	-1	1	-1	1
10	1	10	1	1	-1	-1	-1	-1	1
11	9	11	1	1	-1	-1	-1	1	-1
12	14	12	1	1	1	-1	1	1	-1
13	11	13	1	1	-1	1	-1	1	1
14	16	14	1	1	1	1	1	1	1
15	12	15	1	1	1	1	-1	1	-1
16	3	16	1	1	-1	1	-1	-1	-1

Fig.4 *Design of experiment created in MINITAB*

The input selection parameter is inserted into MATLAB and runs the command. After running the command, the graph of Normal Pattern (N), Up Shift (US), Down Shift (DS), Up Trend (UT), Down Trend (DT) and Cyclic (CYC) appears. The figures below show the graph of the patterns that were recorded from one of the 16 sets, Set 1. After completing the six graphs of the patterns (Fig. 5), the result and data produced will be recalled and loaded to the ANN training command file to get the MSE result. The graph of train data and training progress appears. The MSE value is the performance value.

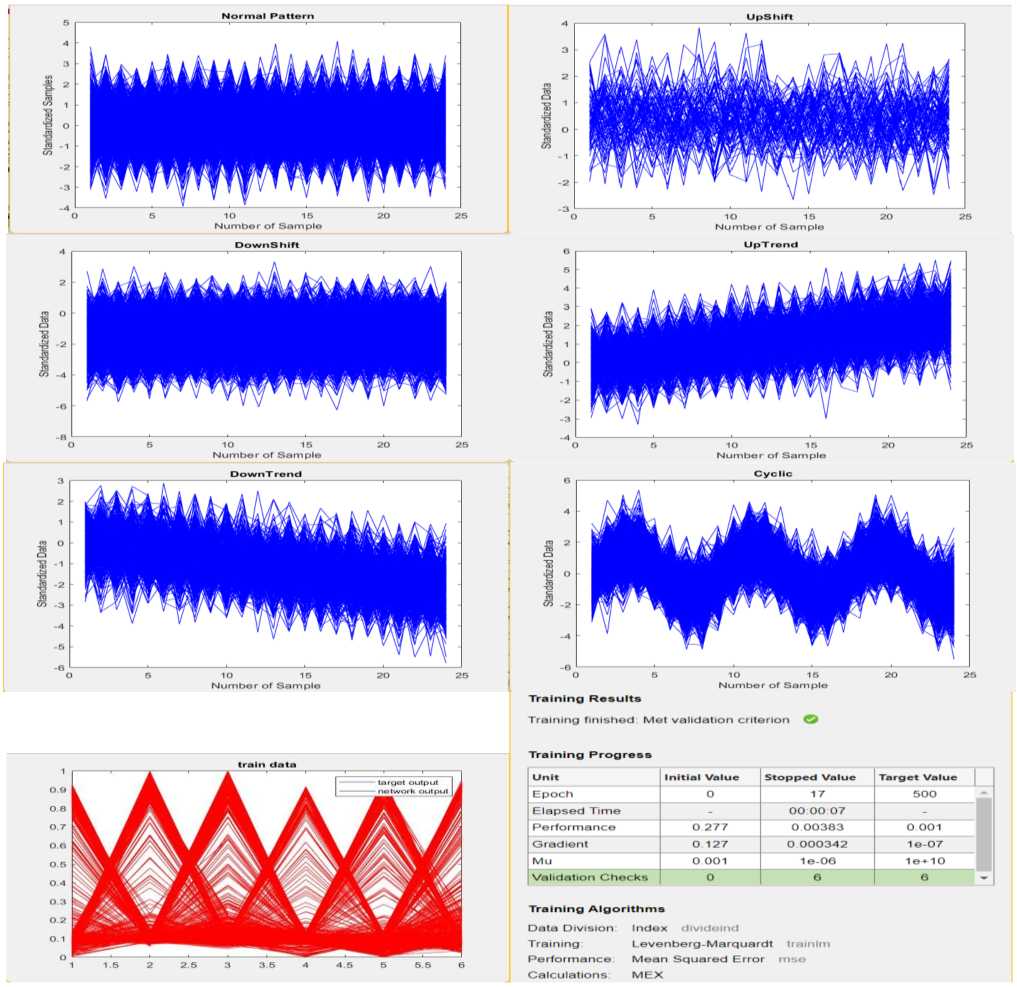

Fig. 5 P*lot graph for six pattern and train data graph*

	%N	%US	%DS	%UT	%DT	%CYC	MSE
1	98.66	99.32	99.25	94.35	90.09	88.29	0.003830
2	100.00	96.80	96.18	96.44	97.22	100.00	0.000992
3	100.00	96.07	95.67	96.63	91.33	100.00	0.003170
4	100.00	92.03	96.03	96.72	94.83	100.00	0.001460
5	100.00	95.26	95.22	94.84	93.10	100.00	0.004760
6	91.46	96.00	96.34	97.51	95.94	95.00	0.005590
7	100.00	95.96	93.58	91.74	90.60	99.10	0.003160
8	100.00	98.87	96.91	91.67	88.33	99.54	0.003790
9	97.22	95.47	97.05	99.02	92.15	91.12	0.003360
10	100.00	91.10	91.03	93.43	97.57	100.00	0.001680
11	100.00	95.49	93.87	90.09	94.69	100.00	0.002800
12	92.73	98.91	97.57	80.65	96.55	95.70	0.004620
13	100.00	93.28	93.33	96.00	95.52	100.00	0.002860
14	100.00	97.65	95.98	95.12	92.20	100.00	0.001640
15	100.00	96.55	98.56	95.15	93.23	100.00	0.001140
16	100.00	93.01	93.94	90.52	94.87	100.00	0.001780

Fig. 6 *Result of running parameter in MATLAB*

The data in Fig. 6 was inserted into the MINITAB to calculate the mean data for each percentage (%) and MSE. The mean of main effects plot result is recorded in Fig. 7.

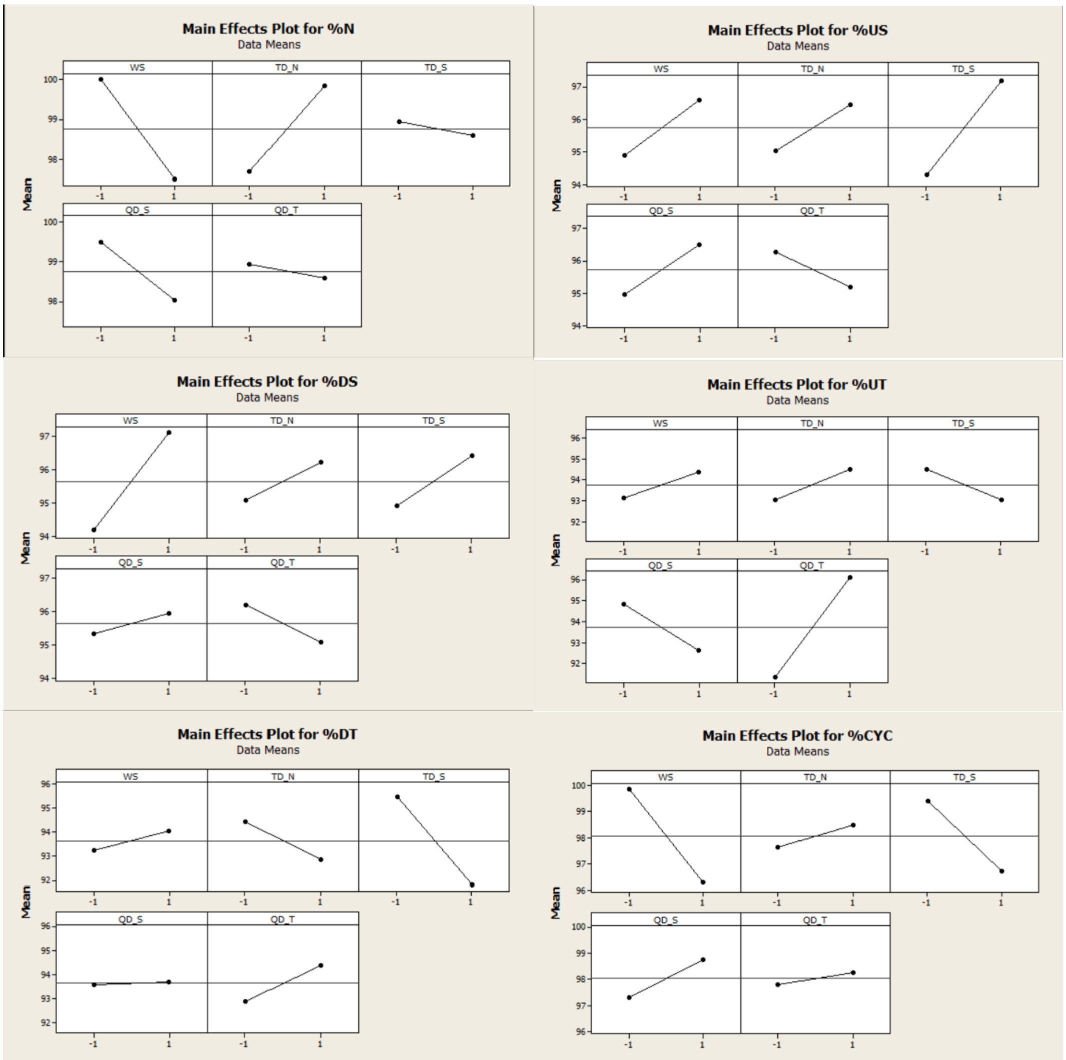

Fig. 7 *Mean of main effects plot for percentage each pattern*

The data of the mean is transferred into the table below. This data later will be used to create another four set of data for running selected from the best, which is the mean.

Table 3 *New set of parameter combination from mean of main effects plot*

	WS	TD_N	TD_S	QD_S	QD_T
New Set 1	1(24)	1(1500)	-1(50)	1(0.1)	-1(0.002)
New Set 2	1(24)	1(1500)	-1(50)	1(0.1)	1(0.001)
New Set 3	-1(20)	1(1500)	-1(50)	1(0.1)	-1(0.002)
New Set 4	-1(20)	1(1500)	-1(50)	1(0.1)	1(0.001)

From this new set data (Table 3), the objective is to obtain the highest value of percentage (%) Normal Pattern (N) and balance with the other abnormal pattern, which is the average percentage (%) value of abnormal pattern. The new set data is rerun in MATLAB to obtain the result. Then the result for this testing has been recorded and summed up as shown in the table.

Table 4 *Result of the simulation of new set in MATLAB*

	MSE	%N	%US	%DS	%UT	%DT	%CYC
New Set 1	0.00152	100.00	97.84	96.24	92.50	89.92	100.00
New Set 2	0.00130	100.00	95.27	95.41	92.61	95.11	100.00
New Set 3	0.00337	100.00	95.80	95.86	96.46	93.22	100.00
New Set 4	0.00323	100.00	90.35	95.22	92.57	93.94	100.00

Table 5 *Comparison of normal percentage with average abnormal percentage*

	MSE	Percentage Normal Pattern	Average Percentage Abnormal Pattern
New Set 1	0.00152	100.00	95.30
New Set 2	0.00130	100.00	95.68
New Set 3	0.00337	100.00	96.27
New Set 4	0.00323	100.00	94.42

From the data of the new set shown in Table 4, the best parameter based on MSE, which is closer to 0.001, is the parameter from New Set 2. However, the better parameter based on the balancing percentage normal pattern compared to the average percentage of abnormal pattern, the New Set 3 is the highest recognition accuracy, the average of all percentages is above target, which is 95%, except for the percentage downtrend pattern, which is 93.22%. This result also shows that a low window size is enough to achieve 100% for normal pattern and cyclic patterns. From the New Set 1, it can be concluded that the factor is superior for recognizing shift patterns but has a slightly impaired capability for trend patterns, also known as imbalance recognition.

4. Conclusion

The conclusions of this study are as follows:

- The first objective of this study, which is to develop a basic control chart pattern recognition framework for identifying abnormal control chart patterns, has been successfully achieved. This result can be proved by the graph plot of every parameter running through the command in MATLAB.
- The study's second objective is to select a proper setting parameter for the pattern classifier to improve the classification accuracy. The result of the partial fractional using DOE in MINITAB proves that this study's conclusion of the accuracy is obtained by exceeding the 95% target.

References

[1] Zaman, M., & Hassan, A. (2021, March). X-barControl Chart Patterns Identification Using Nelson's Run Rules. In Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management, Singapore (pp. 7-11).

[2] Masood, I., & Hassan, A. (2013). Pattern recognition for bivariate process mean shifts using feature-based artificial neural network. The International Journal of Advanced Manufacturing Technology, 66, 1201-1218., doi: 10.1007/s00170-012-4399-2.

[3] Masood, I., & Hassan, A. (2012). An integrated MEWMA-ANN scheme towards balanced monitoring and accurate diagnosis of bivariate process mean shifts. Journal of King Saud University-Computer and Information Sciences, 24(2), 93-100., doi: 10.1016/j.jksuci.2011.10.002.

[4] Gauri, S. K., & Chakraborty, S. (2007). A study on the various features for effective control chart pattern recognition. The International Journal of Advanced Manufacturing Technology, 34, 385-398., doi: 10.1007/s00170-006-0591-6.

[5] Addeh, A., Khormali, A., & Golilarz, N. A. (2018). Control chart pattern recognition using RBF neural network with new training algorithm and practical features. ISA transactions, 79, 202-216., doi: 10.1016/j.isatra.2018.04.020.

[6] Rakitzis, A., & Bersimis, S. (2006). Detection of non-random patterns in shewhart control charts: methods and applications. Available at SSRN 944166.

[7] Sağiroğlu, Ş. E. R. E. F., BEŞDOK, E., & ERLER, M. (2000). Control chart pattern recognition using artificial neural networks. Turkish Journal of Electrical Engineering and Computer Sciences, 8(2).

[8] Garcia, E., Penabaena-Niebles, R., Jubiz-Diaz, M., & Perez-Tafur, A. (2022). Concurrent control chart pattern recognition: A systematic review. Mathematics, 10(6), 934, doi: 10.3390/math10060934.

[9] Noskievičová, D. (2013). Complex control chart interpretation. International Journal of Engineering Business Management, 5, 13., doi: 10.5772/DOI.

[10] Chowdhury, N. R., & Janan, F. (2020). Control chart pattern recognition: A comparison between statistical correlation measure and support vector machine (svm). In Proceedings of the 5th NA International Conference on Industrial Engineering and Operations Management Detroit (pp. 10-14).

[11] Masood, I., Abidin, N. Z. Z., Roshidi, N. R., Rejab, N. A., & Johari, M. F. (2013). Design of an Artificial Neural Network Pattern Recognition Scheme Using Full Factorial Experiment. In Applied Mechanics and Materials (Vols. 465–466, pp. 1149–1154). Trans Tech Publications, Ltd. https://doi.org/10.4028/www.scientific.net/amm.465-466.1149

[12] Hassan, A. (2011). An improved scheme for online recognition of control chart patterns. International Journal of Computer Aided Engineering and Technology, 3(3-4), 309-321.