**Indonesian Journal of Electrical Engineering and Computer Science**
Vol. 31, No. 2, August 2023, pp. 734~746
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v31.i2.pp734-746 

# An improved features selection approach for control chart patterns recognition

**Waseem Alwan[1], Nor Hasrul Akhmal Ngadiman[1], Adnan Hassan[1], Mohd Syahril Ramadhan Mohd Saufi[1], Azanizawati Ma'aram[1], Ibrahim Masood[2]**
[1]Faculty of Mechanical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia
[2]Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia, Batu Pahat, Malaysia

**Article Info**

***Article history:***

Received Nov 23, 2022
Revised Mar 21, 2023
Accepted Apr 2, 2023

***Keywords:***

Artificial neural network
Control chart patterns
Correlation coefficient
Feature selection algorithm
Relief algorithm

**ABSTRACT**

Control chart patterns (CCPs) are an essential diagnostic tool for process monitoring using statistical process control (SPC). CCPs are widely used to improve production quality in many engineering applications. The principle is to recognize the state of a process, either a stable process or a deterioration to an unstable process. It is used to significantly narrow the set of possible assignable causes by shortening the diagnostic process to improve the quality. Machine learning techniques have been widely used in CCPs. Artificial neural networks with multilayer perceptron (ANN-MLP) are one of the standard tools used for this purpose. This paper proposes an improved features selection method to select the best features as input representation for control chart patterns recognition. The results demonstrate that the proposed approach can effectively recognize CCPs even for small patterns with a mean shift of less than 1.5 sigma. The dimensional reduction was achieved by employing Relief, correlation, and Fisher algorithms (RCF) for feature selection and (ANN-MLP) as a classifier (RCF-ANN). This study provides an experimental result that compares the performance before and after dimensional reduction.



***Corresponding Author:***

Waseem Alwan
Faculty of Mechanical Engineering, Universiti Teknologi Malaysia
81310 Skudai, Johor Bahru, Johor, Malaysia
Email: zaboon.w@graduate.utm.my

## 1. INTRODUCTION

As an essential tool for statistical process control (SPC), the control chart plays a vital role in the quality control of various production processes. The principle of control chart pattern recognition is to identify the state of process if working under the standard condition and be stable under the quality characteristic, or the process deteriorates from stable to unstable. There are many research aspects involving CCPs to improve production quality [1]. The control chart patterns (CCPs) with artificial neural networks (ANNs) can recognize the quality in many sectors, such as the health, food industry, and financial market [2]. Many unstable patterns exist, but most previous studies focused on the five common abnormal types studied in addition to the normal pattern, its cycle, increase trend, decrease trend, upper shift, and downward shift [3]-[8] among others.

The input data representation depends on the dimensional of data, even its raw data or features-based. The features-based has better accuracy than raw data and improve performance by removing redundant and irrelevant features [9], [10]. Many features which extracted from raw data from the previous studies' investigation such as wavelet features [11], shape features [7], and statistical features [5]. Some researchers used mixed (statistical and shape) features to improve the performance [12]. They reported good accuracy when using mixed features [13], [14].

Many feature selection algorithms are being applied to get good data representation and remove the reiterated features like a genetic algorithm (GA) [13], principal component analysis (PCA) [14], and kernel entropy component analysis (KECA) as a feature reduction algorithm [15]. Feature selection plays a critical role in monitoring the control charts and pattern recognition, driven by decreasing input dimensionality in target problems and growing interest in advanced.

The Relief algorithm are a distinctive family of filter-style feature selection algorithms that have gained popularity by finding an appropriate balance between these goals while reacting flexibly to diverse data variables. The feature selection process reduces the dimensionality of data sets in classification and regression tasks. It eliminates redundant and irrelevant characteristics that have little influence on the model's aim. This approach can minimize computing costs and training time, reduce overfitting, reduce complexity, and enhance acceptable accuracy based on minimum features. The Relief method is a common and straightforward instance-based learning technique [16], [17]. The core idea behind RAs is to estimate the quality of features based on how well the attribute can distinguish between instances that are near to each other. The correlation coefficient quantifies the degree of the linear relationship between variables, and has an exact value less than or equal to one. The larger the degree of linear correlation between the two variables, the higher the exact value of the correlation coefficient is to 1. As a result, the absolute value of the correlation coefficient between features should be used to assess the significance of duplicated features. Utilizing the correlation coefficient's absolute value as the weight of a redundant element, use the minimization principle in the assessment criteria to choose the feature subset with the most classification information, enhancing classification accuracy [18]. The Fisher score is a feature selection method, it is used to reduce dimensions by selecting the best features. With selected m features, $X \in R_{dxn}$ input data matrix reduces to $Z \in R_{mxn}$ [19]. The top m-ranked features are selected by calculated Fisher score for each feature. The reduced dataset will be the sub-dataset containing the most significant m features in the original dataset consisting of m features [20]. The Fisher Z transformation is a formula we can use to transform Pearson’s correlation coefficient (r) into a value (zr) that can be used to calculate a confidence interval for Pearson's correlation coefficient. Fisher's Z calculates confidence intervals for both r and correlation differences. However, it is typically employed to examine the significance of a difference in two correlation coefficients, r1, and r2, from sample data [21].

Most previous studies depend in their studies on one feature selection algorithm. So, in this paper, we proposed an improved approach to select the best feature sets extracted from raw data by employing three feature selection algorithms (Relief, correlation, and Fisher) with an ANN-based classifier to detect a small change (±1.5 sigma) in the unstable process. The significance of this strategy is that it strengthens feature selection decision making by relying on three selection algorithms rather than individual selection algorithms to reduce wrong selection of redundant features. The paper is organized into four sections. Section 2 provides the methodology; section 3 presents the results and discussion, and section 4 concludes the paper.

## 2. METHOD

### 2.1. Data generation

In this paper, we focus on six common patterns that studies in previous studies (normal, cycle, increase trend, decrease trend, upper shift, and downward shift) as in the previous studies [5], [11], [22]. Using Monti Carlo simulation rather than real data because it’s not economy as previous studies to generate all type of patterns. A total of 6,000 X-bar chart patterns were generated (1,000 patterns for each type) using (1) to (6).

$$Normal\ (NOR) y\_t = \mu + r\_i\ \sigma \tag{1}$$

$$Cyclic\ (CYC)\ y\_t = \ \mu + r\_i\ \sigma + a\ [\![ sin(2\pi i\ /\ T ]\!]\ ) \tag{2}$$

$$Increase\ trend\ (IT)\ y_t = \mu + r_i\sigma + gi \tag{3}$$

$$Decrease\ trend\ (DT)\ y_t = \mu + r_i\sigma - gi \tag{4}$$

$$Upward\ shift\ (US)\ y_t = \mu + r_i\sigma \mp ks \tag{5}$$

$$Downward\ shift\ (DS)\ y_t = \mu + r_i\sigma \mp ks \tag{6}$$

The parameters for the equations which used in previous studies are given in Table 1. In this study the authors carefully selected the parameters to be suitable with previous research as shown in Table 2.

Table 1. The parameters and values of the equations utilized in previous literature

| Ref. Pram. | Normal Mean Std | Stratification Random noise | Systematic Systematic departure | Cyclic Amplitude Period | Increase Trend Gradient | Decrease Trend Gradient | Upward Shift magnitude position | Downward Shift magnitude position | RA (%) |
|---|---|---|---|---|---|---|---|---|---|
| [9] | $\mu$=0 $\sigma$=1 | $1/3(\sigma)$ | | $a = 0.5\sigma – 2.5\sigma$ $p = 10$ | $0.015\sigma – 0.025\sigma$ | $(-0.025\ -0.015)\sigma$ | $0.7\sigma$ to $2.5\sigma$ | $-2.5\sigma - -0.7\sigma$ | 95.2 |
| [23] | $\mu$=80 $\sigma$=5 | $0.2\sigma \leq \sigma \leq 0.2\sigma$ | $1\sigma \leq g \leq 3\sigma$ | $1.5\sigma \leq a \leq 2.5\sigma$ $8 \leq p \leq 16$ | $0.05\sigma \leq g \leq 0.1\sigma$ | $-0.1\sigma \leq g \leq -0.05\sigma$ | $1.5\sigma \leq s \leq 2.5\sigma$ $15 \leq P \leq 45$ | $-2.5\sigma \leq s \leq -1.5\sigma$ $15 \leq P \leq 45$ | 99.6 |
| [24] | $\mu$=0 $\sigma$=1 | | $1\sigma \leq g \leq 3\sigma$ | $1.5\sigma \leq a \leq 2.5\sigma$ $4 \leq p \leq 8$ | $0.05\sigma \leq g \leq 0.1\sigma$ | $-0.1\sigma \leq g \leq -0.05\sigma$ | $1.5\sigma \leq s \leq 2.5\sigma$ $11 \leq p \leq 21$ | $-2.5\sigma \leq s \leq -1.5$ $11 \leq p \leq 21$ | <99 |
| [3] | $\mu$=30 $\sigma$=0.05 | | | $1.5\sigma \leq a \leq 4\sigma$ $4 \leq p \leq 8$ | $[0.1\sigma, 0.3\sigma]$ | $[-0.3\sigma, -0.1\sigma]$ | $[1.5\sigma, 3\sigma]$, p= [4-9] | $1.5\sigma, 3\sigma]$, p= [4-9] | 99.3 |
| [8] | $\mu$=0 $\sigma$=1 | $(0.1–0.4)\sigma$ | $(0.005–2.5)\sigma$ | $0 \leq a \leq 1.8$ $p = 10$ | 0.015–0.025 | 0.015–0.025 | s=0.6–2.5 | s=0.6–2.5 | 99.8 |
| [25] | $\mu$=0 $\sigma$=1 | | | $a = [1,3]\sigma$ $p = 8$ | $[0.1,0.3]\sigma$ | $[-0.3- -0.1]\sigma$ | $s = [1\sigma, 3\sigma]$ P= 12,13,14 | $s = [-3\sigma, -1\sigma]$ p= 12,13,14 | 98 |

Table 2. The parameters and values utilized for the six CCPs

| Parameters | Definition | Value |
|---|---|---|
| $\mu$ | Mean. | 0 |
| $\sigma$ | Standard deviation. | 1 |
| $\sigma'$ | Random noise all for each abnormal pattern. | $\sigma' = 1/3\sigma$ |
| $a$ | Amplitude. | $0.5\sigma \leq a \leq 2.5\sigma$ |
| $T$ | Period of cycle. | 8, 10 |
| $s$ | Shift magnitude. | normal shift $1.5\sigma \leq s \leq 2.8\sigma$, small shift $s < 1.5\sigma$ |
| $k$ | Shift position. | *position* = (5,15,20), $k = 1$ if $i \geq$ position else $k = 0$ |
| $g$ | Gradient for a trend pattern. | $0.015\sigma \leq d \leq 0.025\sigma$ |
| $ri$ | At the ith time point, a random. value of a standard normal variate | $-3 \leq r \leq +3$ |
| $yi$ | Time series value at ith time point | 1-30 |
| Standardized: N (0,1) | | |

### 2.2. Features extraction

The feature extraction from raw data decreases the dimensional input for machine learning, then increases the recognition efficiency when making the network size small [15]. Two common features are working correctly with CCPs statistical features and shape features. We focus on the common 13 mixed (statistical and Shape) features used with CCPs [12].

#### 2.2.1. Statistical features

The 10 candidate statistical features examined in this research (mean, standard deviation, maximum value, minimum value, skewness, Kurtosis, mean square error, slop, mean square value, and cumulative summation) as shown in Figure 1. Its chosen based on previous studies [8], [10], [26]. The mean feature for normal and cycle is approximately equal to zero when for increase trend and upper shift have value more than zero. Moreover, decrease trend and down shift have a value less than zero. This feature can differentiate the normal pattern and cycle from other patterns, as shown in Figure 1(a). The standard deviation feature can differentiate between normal and other patterns, as shown in Figure 1(b). The maximum value is higher for the increase trend, and upper shift pattern, less for normal and cycle, and minimum for the decrease trend and down shift, as shown in Figure 1(c). The minimum value is higher for the decrease trend and down shift pattern and less than for normal and cycle and minimum for increase trend and upper shift as shown in Figure 1(d). Skewness is higher for trend and normal and minimum for cycle and shift, as shown in Figure 1(e). Kurtosis is higher for increase trend and upper shift from other patterns, as shown in Figure 1(f). The mean square value is minimum for normal and higher for other patterns, as shown in Figure 1(g). The slope has a higher positive value for the increase trend and upper shift, approximately equal to zero for normal and cycle, and a higher negative for the decrease trend and down shift, as shown in in Figure 1(h). The mean square error is higher for shift and cycle than the other patterns shown in in Figure 1(i). Cumulative summation has a higher value for increase trend and upper shift and less than for cycle and minimum for normal and decrease trend and down shift as shown in in Figure 1(j).

#### 2.2.2. Shape features

This study selected three typical shape features (Least square slop, APML, and APLS) [27] as shown in Figure 2. The least-square slope has a higher value for trend and shift and a minimum for normal and cycle, as shown in Figure 2(a). APML has a higher value with trend and shift and minimum value for normal and cycle, as shown in Figure 2(b). APLS has a higher positive value for increase trend and upper shift and minimum value for normal and cycle and a higher negative value for decrease trend and down shift pattern, as shown in Figure 2(c).

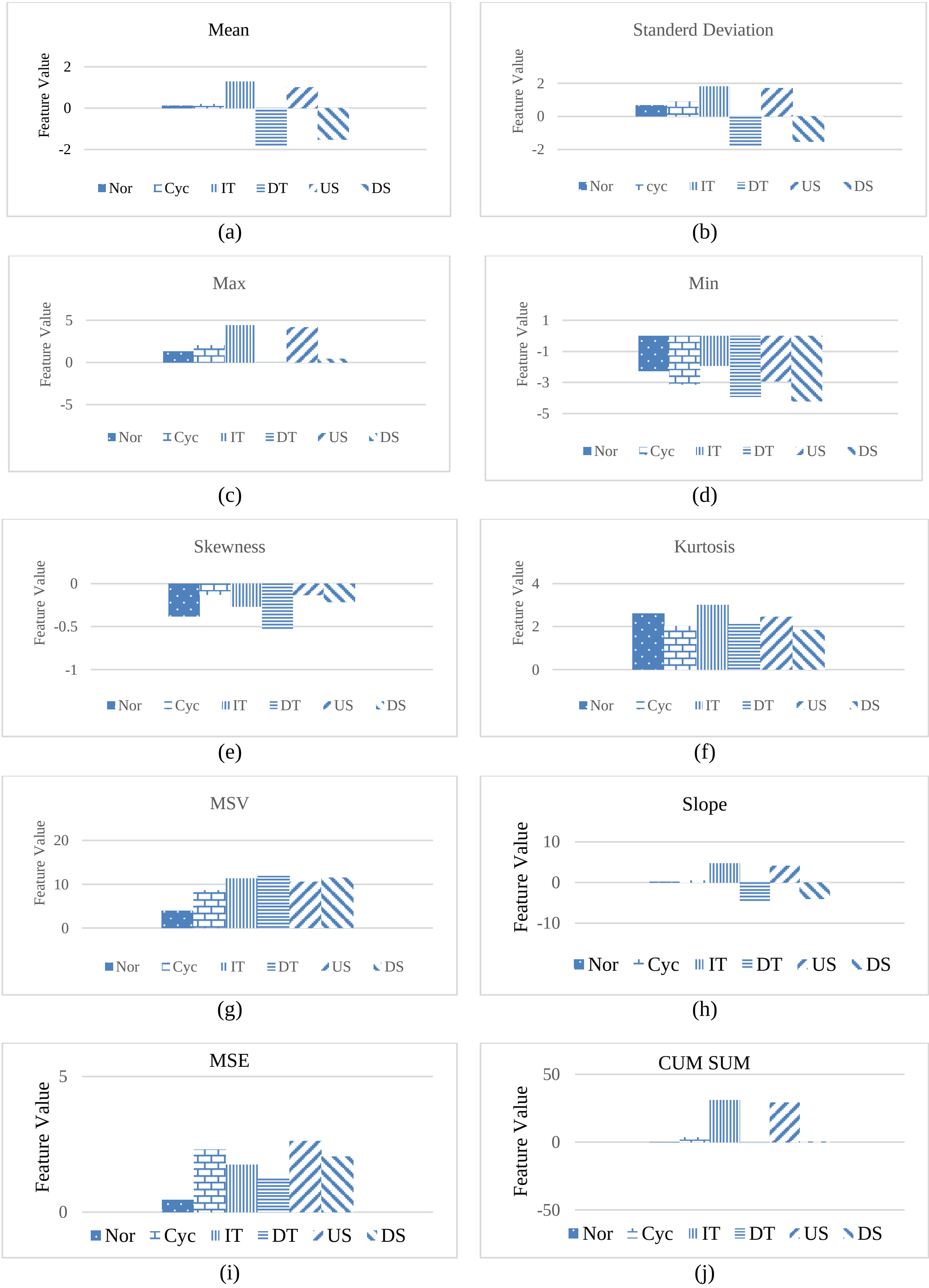
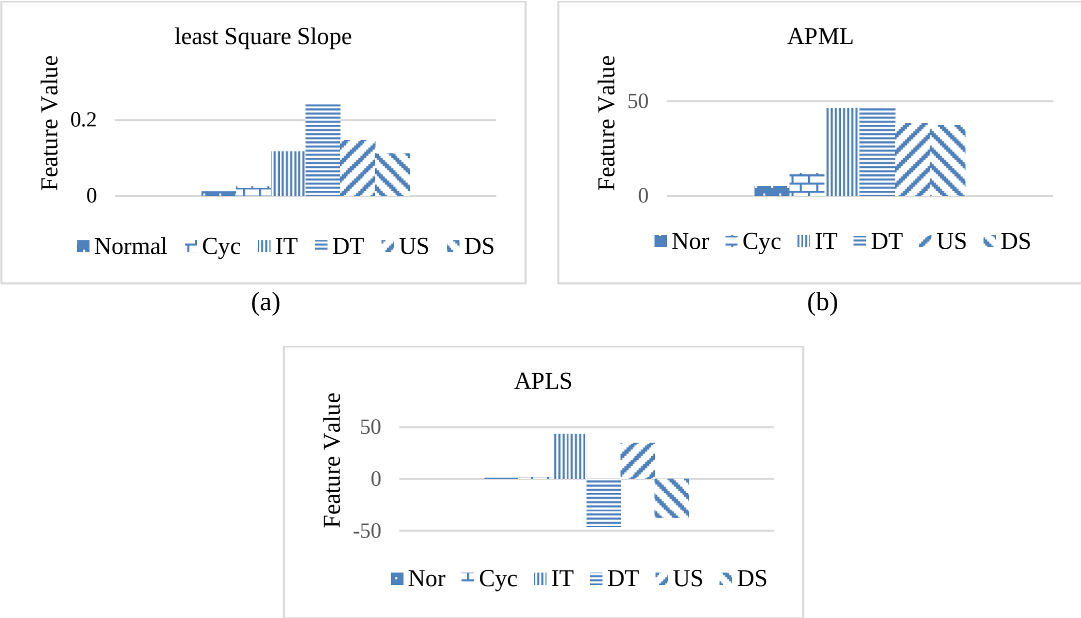



Figure 1. Statistical features for each pattern (a) mean, (b) stander deviation, (c) maximum, (d) minimum, (e) skewness, (f) Kurtosis, (g) mean square value, (h) slope, (i) mean square error, and (j) cumulative summation




(a) (b) (c)

Figure 2. Shape features for each pattern (a) least-square slope, (b) APML, and (c) APLS

### 2.3. Features selection

Machine learning approaches were utilized to build control chart pattern recognition (CCPR) models that differentiate and identify different patterns class. These models' accuracy is determined by the features used and the quality of information collected. In general, not all of the features utilized in pattern recognition are efficient at distinguishing the patterns. Some features could be redundant or useless and maybe lower the accuracy of the classifier. As a result, feature selection is critical in implementing machine learning techniques [12]. For that, filtration needs to make the features as soon as possible and represent the data. For this purpose, we employ three feature algorithms to select the vital feature that can present the type of pattern by minimum features.

These three algorithms were used in previous studies on many felids. However, this first time will use together with CCPs as a new approach to selected features it is (relief algorithm, correlation coefficient, and fisher coefficient) [28], [29]. Then select the best six features in all three algorithms and depend on them as an input for the network.

### 2.4. Pattern recognizer design

Many classification approaches are available in the literature, such as ANN, SVM, and decision tree (DT) [30]. The multilayer perceptron's (MLPs) architecture was employed as a recognizer as it has been applied to solve more complex problems, such as prediction and modelling in CCPs [7], [31]-[33]. It has three layers; the first layer is the input layer representing the input in this paper, corresponding to the number of features (6). The second layer is called the hidden layer; its sets one hidden layer; the number of nodes in the hidden layer is set empirically equal to 12. The third layer is the output layer corresponding to the number of patterns we study equally (6). Then the network (6×12×6). Levenberg-Marquardt (trainlm) algorithm was employed as a learning algorithm after testing other learning algorithms. Broyden–Fletcher–Goldfarb–Shanno (BFGS) quasi-Newton gradient descent with momentum and adjustable learning rate (traingdx), (trainbfg).

The first step is data generation of 6,000 patterns, 1,000 patterns for each pattern depending on (1) to (6), extraction of the 13 mix-features from raw data then select the best 6 features depending on three algorithms results (Relief, correlation, and Fisher). After selecting features, we must make (normalization) to those features between (-1, +1) using (7). This process is better to make the network work well and get the generalization. During the training phase, the data must be labeled for each class of patterns, the targeted values for the recognizers' output nodes for the correct class will label as 0.9, and another wrong class will be labeled as 0.1, as shown in Table 3. This dataset was divided into training (70%), validation (10%), and preliminary testing (20%) before the sample data were presented to the ANN for the learning process. 4,200 patterns for training and used for updating the network weights and biases. 600 patterns were used for validation, and 1,200 patterns were tested, and these patterns like hidden during the training phase. The parameters and training specifications of the network are set as the number of epochs between showing the progress.

$$P_n = \frac{2(P - P_m)}{(P_{max} - P_{min})} - 1 \quad (7)$$

Where: $P$=observed feature value. $P_n$=normalized observed feature value. $P_{min}$=minimum feature value for the feature. $P_{max}$=maximum feature value for the feature.

Table 3. Recognizer output that are targeted [9]

| Pattern's class | Definition | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| 1 | NOR | 0.9 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 |
| 2 | CYC | 0.1 | 0.9 | 0.1 | 0.1 | 0.1 | 0.1 |
| 3 | IT | 0.1 | 0.1 | 0.9 | 0.1 | 0.1 | 0.1 |
| 4 | DT | 0.1 | 0.1 | 0.1 | 0.9 | 0.1 | 0.1 |
| 5 | US | 0.1 | 0.1 | 0.1 | 0.1 | 0.9 | 0.1 |
| 6 | DS | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.9 |

The flow chart for the training phase and testing is shown in Figure 3. Figure 3(a) show the training phase and Figure 3(b) show the testing phase, respectively. The maximum number of epochs=500. The learning rate is set as 0.5. Momentum constant set as 0.5. Performance measurement (MSE). Performance goal set as 10-3. All the procedures were coded in MATLAB R2017a using its ANN toolbox.

To evaluate the efficiency of the RCF-ANN approach, we test the network performance with 13 common extraction features that extraction from raw data as input and the six features after applying RCF and comparing the results. The confusion matrix is the evaluation way to see the correct recognition.

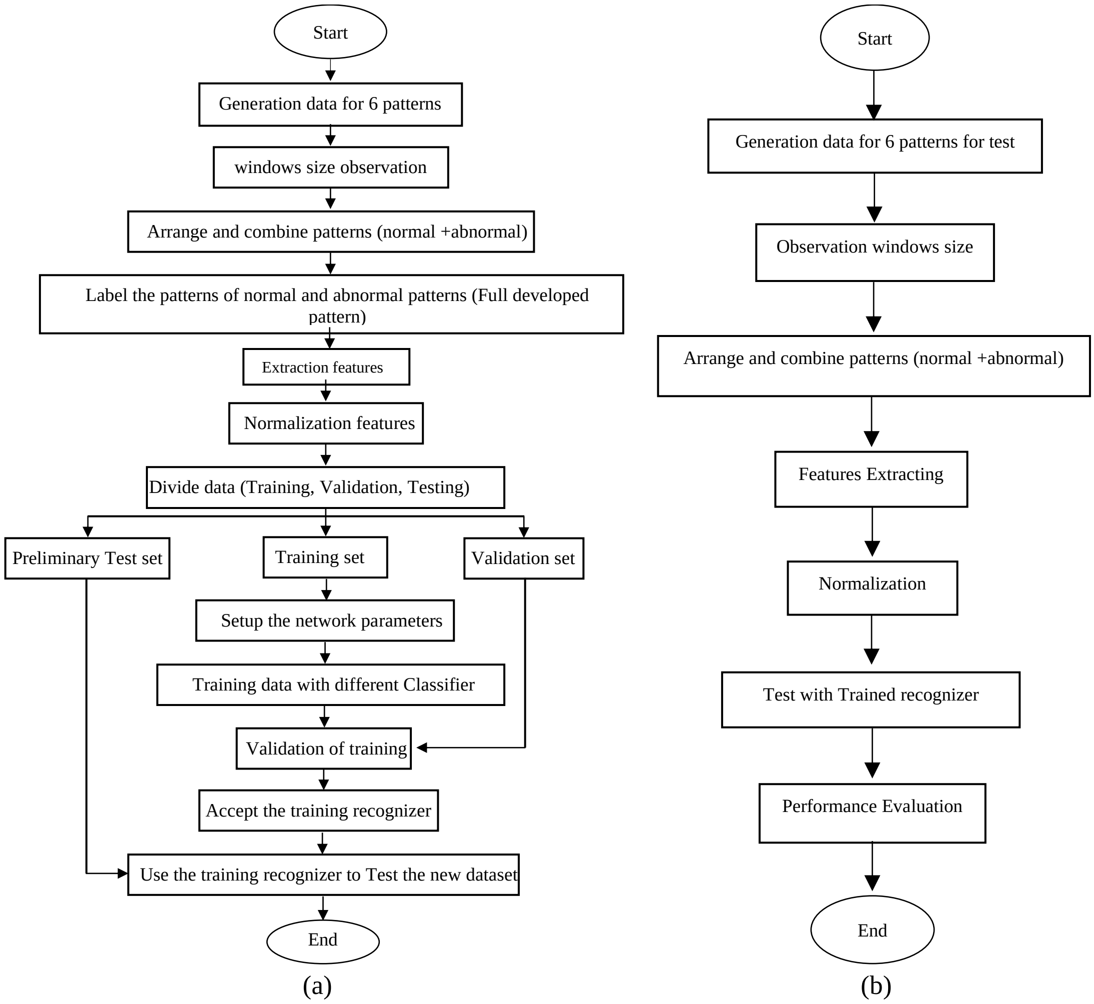
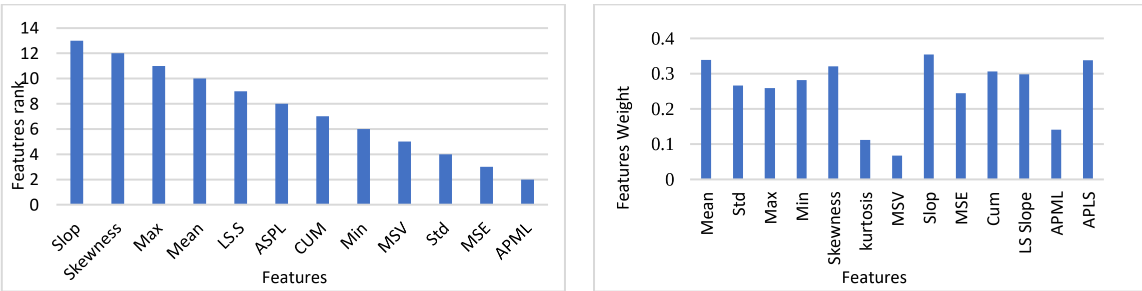



Figure 3. The study flow chart for (a) training and (b) testing

## 3. RESULTS AND DISCUSSION
### 3.1. Results of relief algorithm

Kira and Rendell Kononenko developed the Relief algorithm, which is based on instant-based learning, and was later improved by [34]. Relief algorithm is quite effective in estimating features. It chooses the traits that are relevant to the goal. It finds the nearest neighbor by randomly selecting a case from the data. The feature values of the nearest neighbors are compared to the sampled instance to update the relevance score for each feature. Relief looks for two nearest neighbors for a given instance: one from the same class (nearest hit) and one from a different class (called nearest miss) [34]. Relief's estimate W [A] of feature (A) is an approximation of the following probability difference: W [A]=P (different value of the closest instance from a different class)-P (different value of the closest instance from the same class). The reasoning is that a good characteristic should distinguish between examples of various classes while having the same value for instances of the same class [28]. This features selection algorithm calculates two factors the rank of features and the weight as shown in Figure 4. Most previous researchers depend on the weight because it gives the ideal presentation of the nearest features from the target.

In this paper, we depend on weight to select the best feature. There are two parameters in this algorithm that must to known. In this study, the initial weight for all features=0.2 [28]. Moreover, the number of valid features equals six. Figure 4(a) shows the output of the ranked features in the Relief algorithm. The algorithm arranged all the features depending on rank. In this experimental study, the features slope ranks in the height position, and its weight is 0.2961. the weight value of 0.2069 represents the median value of the Relief results, as shown in Figure 4(b). For that, the features with a weight above the 0.2069 value are considered the selected features. Hence the features slope, maximum, mean, least square slope (LS.S), cumulative summation (Cum.Sum), minimum (Min), standard deviation (St.D), mean square error (MSE), skewness, and area between the pattern and its least-square line (APLS) are selected as independent features in the output of the relief algorithm. In order to ensure the features, the next step is to apply the correlation algorithm.



(a) (b)

Figure 4. Relief features results depend on (a) rank and (b) weight

### 3.2. Results of correlation coefficient

The correlation coefficient measures the strength of the association between the features statistically. The values vary from (-1.0 to 1.0). A computed value greater than 1.0 or less than -1.0 indicates that the correlation measurement was incorrect. A correlation of (-1.0) indicates a perfect negative correlation, whereas one of 1.0 indicates a perfect positive correlation. A correlation of 0.0 indicates no linear link between the two variables' movements. Table 4 represents the correlation coefficient results.

Table 4. Correlation coefficient results

| Feature | Slope | Skw. | Max | Mean | LS. S | Kurt | Cum | Min | MSV | Std | MSE | APML | APSL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Slope | 1.0 | 0.0 | 0.9 | 0.9 | 0.5 | -0.9 | -0.5 | 1.0 | 0.0 | 0.9 | 1.0 | -0.6 | 1.0 |
| Skw. | 0.0 | 1.0 | 0.4 | -0.4 | 0.1 | 0.3 | 0.7 | -0.1 | 1.0 | 0.2 | -0.1 | 0.3 | 0.0 |
| Max | 0.9 | 0.4 | 1.0 | 0.6 | 0.6 | -0.6 | 0.0 | 0.9 | -0.4 | 0.9 | 0.8 | -0.3 | 0.9 |
| Mean | 0.9 | -0.4 | 0.6 | 1.0 | 0.5 | -0.9 | -0.7 | 0.9 | -0.4 | 0.7 | 0.9 | -0.7 | 0.9 |
| LS. S | 0.5 | 0.1 | 0.6 | 0.5 | 1.0 | -0.4 | 0.1 | 0.6 | 0.0 | 0.7 | 0.7 | 0.2 | 0.5 |
| Kurt | -0.9 | 0.3 | -0.6 | -0.9 | -0.4 | 1.0 | 0.8 | -0.9 | 0.2 | -0.7 | -0.9 | 0.8 | -0.9 |
| Cum | -0.5 | 0.7 | 0.0 | -0.7 | 0.1 | 0.8 | 1.0 | -0.5 | 0.6 | -0.1 | -0.5 | 0.9 | -0.5 |
| Min | 1.0 | -0.1 | 0.9 | 0.9 | 0.6 | -0.9 | -0.5 | 1.0 | -0.1 | 0.9 | 1.0 | -0.5 | 1.0 |
| MSV | 0.0 | 1.0 | 0.4 | -0.4 | 0.0 | 0.2 | 0.6 | -0.1 | 1.0 | 0.1 | -0.1 | 0.2 | 0.0 |
| Std | 0.9 | 0.2 | 0.9 | 0.7 | 0.7 | -0.7 | -0.1 | 0.9 | 0.1 | 1.0 | 0.9 | -0.2 | 0.9 |
| MSE | 1.0 | -0.1 | 0.8 | 0.9 | 0.7 | -0.9 | -0.5 | 1.0 | -0.1 | 0.9 | 1.0 | -0.5 | 1.0 |
| APML | -0.6 | 0.3 | -0.3 | -0.7 | 0.2 | 0.8 | 0.9 | -0.5 | 0.2 | -0.2 | -0.5 | 1.0 | -0.6 |
| APLS | 1.0 | 0.0 | 0.9 | 0.9 | 0.5 | -0.9 | -0.5 | 1.0 | 0.0 | 0.9 | 1.0 | -0.6 | 1.0 |

From Table 4 the results showed the correlation among all the features. In the first row show the correlation between the slope and Max, Mean, Min, Std, MSE, and APLS. We selected the higher correlation between the features and neglected the lower correlation. The threshold set as 0.6, as shown in Table 5.

Table 5. Best correlation between the features

| Feature | Slope | Max | Mean | LS. S | Min | Std | MES | APLS |
|---|---|---|---|---|---|---|---|---|
| Slop | 1.0 | 0.9 | 0.9 | | 1.0 | 0.9 | 1.0 | 1.0 |
| Max | | 1.0 | 0.6 | 0.6 | 0.9 | 0.9 | 0.8 | 0.9 |
| Mean | | | 1.0 | | 0.9 | 0.7 | 0.9 | 0.9 |
| LS. S | | | | 1.0 | 0.6 | 0.7 | 0.7 | |
| Min | | | | | 1.0 | 0.9 | 1.0 | 1.0 |
| Std | | | | | | 1.0 | | 0.9 |
| MSE | | | | | | | 1.0 | 1.0 |
| APLS | | | | | | | | 1.0 |

The correlation coefficient of more than 0.6 indicates a strong positive correlation between the features. There might be a causal relationship between the two correlated variables. Furthermore, if there is a link, it may be indirect. A portion of the variation in one of the variables, as measured by its variance, may be attributed to its link with the causes of the other variables.

### 3.3. Results of fisher coefficient

By applying Fisher's exact test to our scenario, the Fisher Z transformation formula used in this paper would compute the following exact probability. The Fisher is necessary in order to enable a comparison of the means. The results of Fisher are shown in Table 6.

Table 6. Fisher coefficient results

| Feature | Slope | Skew | Max | Mean | LS. S | Kur | Cum | Min | MSV | Std | MSE | APML | APLS |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Slope | Inf | -0.335 | 0.819 | 1.6 | -0.99 | -0.41 | 0.301 | 2.901 | -0.39 | 1.401 | 1.985 | 0.412 | 5.84 |
| Skew | -0.34 | Inf | 0.473 | -0.86 | -0.31 | -0.79 | 0.742 | -0.23 | 2.911 | -0.1 | -0.456 | 0.304 | -0.33 |
| Max | 0.819 | 0.473 | Inf | 0.313 | -1.24 | -1.01 | 0.92 | 1.41 | 0.404 | 1.751 | 1.221 | 0.605 | 1.34 |
| Mean | 1.379 | -0.86 | 0.313 | Inf | -0.39 | -0 | -0.18 | 1.643 | -0.91 | 0.788 | 1.615 | 0 | 1.598 |
| LS. S | -0.99 | -0.305 | -1.24 | -0.389 | Inf | 1.138 | -1.02 | -1.26 | -0.24 | -1.42 | -0.881 | -0.93 | -0.99 |
| Kurt | -0.41 | -0.793 | -1.01 | -0.004 | 1.137 | Inf | -0.76 | -0.49 | -0.77 | -0.52 | -0.28 | 1.1 | -0.41 |
| Cum | 0.301 | 0.742 | 0.92 | -0.181 | -1.02 | -0.76 | Inf | 0.495 | 0.625 | 0.717 | 0.242 | 1.341 | 0.303 |
| Min | 2.147 | -0.229 | 0.858 | 1.043 | -1.26 | -0.49 | 0.495 | Inf | -0.3 | 2.195 | 2.401 | 0.652 | 2.890 |
| MSV | -0.39 | 2.911 | 0.404 | -0.905 | -0.24 | -0.77 | 0.625 | -0.3 | Inf | -0.18 | -0.531 | 0.212 | -0.39 |
| Std | 1.582 | -0.104 | 0.966 | 0.788 | -1.42 | -0.53 | 0.717 | 2.195 | -0.18 | Inf | 1.34 | 0.909 | 1.391 |
| MSE | 1.985 | -0.456 | 0.596 | 1.375 | -0.88 | -0.29 | 0.242 | 1.981 | -0.53 | 1.413 | Inf | 0.42 | 1.997 |
| APML | 0.413 | 0.304 | 0.605 | 2E-04 | -0.94 | -0.43 | 1.541 | 0.652 | 0.212 | 0.909 | 0.424 | Inf | 0.416 |
| APLS | 5.843 | -0.337 | 0.815 | 1.378 | -0.99 | -0.41 | 0.303 | 2.165 | -0.4 | 1.591 | 1.997 | 0.416 | Inf |

From Table 6, the Fisher coefficient for slope is inf. Moreover, other features have negative and positive values. The threshold of this feature value by Selected the positive values of other features and neglected the negative values. The results observed the best Fisher coefficient among the features as in Table 7.

Table 7. Higher Fisher coefficient results

| Feature | Mean | Min | MSV | Std | MSE | APML | APSL |
|---|---|---|---|---|---|---|---|
| Slope | 1.6 | 2.9 | | 1.4 | 2 | | 5.8 |
| Skew | | | 2.9 | | | | |
| Max | | 1.4 | | 1.7 | 1.2 | | 1.3 |
| Mean | | 1.6 | | | 1.6 | | 1.6 |
| LS. S | | | | | | | |
| Kurt | | | | | | 1.1 | |
| Cum | | | | | | 1.3 | |
| Min | | | | 2.1 | 2.4 | | 2.9 |
| Std | | | | | 1.3 | | 1.4 |
| MSE | | | | | | | 2 |

From the above three feature algorithm results, we focus on the best features with higher values for each algorithm and select them. It is also selected during the other algorithm to reassure all the features selected from each method. The feature (Max) was selected in Relief and correlation algorithms, but it is neglected in the Fisher algorithm. The feature (Cum) was selected just in the Relief algorithm but neglected during the correlation and Fisher algorithm. The summary results of all the feature selection methods are shown in Table 8.

Table 8. Summary results of the three methods

| Feature | Relief | Correlation | Fisher |
|---|---|---|---|
| Slope | Selected | Selected | Selected |
| Skewness | X | X | X |
| Max | Selected | Selected | X |
| Mean | Selected | Selected | Selected |
| LS slope | Selected | Selected | X |
| Kurtosis | X | X | X |
| Cum | Selected | X | X |
| Min | Selected | Selected | Selected |
| MSV | X | X | X |
| Std | Selected | Selected | Selected |
| MSE | Selected | Selected | Selected |
| APML | X | X | X |
| APSL | Selected | Selected | Selected |

Depending on those results above, we consider the best six features already selected through all three algorithms as input for the network (Mean, Std, Slope, Min, MSE, APSL). The result before RCF shows weak recognition accuracy of 97.33% and 95.87% for normal shifting (1.5 σ-2.8 σ) and small shifting less than (1.5 σ), respectively, as shown in Tables 9 and 10. The results obtained after employing feature selection to decrease the dimensionality from 13 to six features show good recognition accuracy. Tables 11 and 12 indicate 98.84% and 98.14% for normal shift (1.5-2.8) σ and small shift less than (1.5) σ, respectively. The correct normal pattern recognition was improved when applied features selection from 98% to 100% for the normal shift dataset. And from 96.35% to 99.9 for the small shift dataset.

Table 9. Normal shift (1.5–2.8) sigma with 13 features

| Pattern | NOR | CYC | IT | DT | US | DS |
|---|---|---|---|---|---|---|
| NOR | 98 | 0 | 1.2 | 0 | 0.8 | 0 |
| CYC | 1.47 | 97.52 | 0 | 0 | 1 | 0 |
| IT | 0 | 0 | 97.76 | 0 | 2.23 | 0 |
| DT | 0 | 0 | 0 | 96.65 | 0 | 3.34 |
| US | 0 | 0 | 2.72 | 0 | 97.27 | 0 |
| DS | 1 | 0 | 0 | 2.19 | 0 | 96.80 |

Table 10. Small shift less than (1.5) sigma with 13 features

| Pattern | NOR | CYC | IT | DT | US | DS |
|---|---|---|---|---|---|---|
| NOR | 96.35 | 0 | 0 | 1.6 | 0 | 2.05 |
| CYC | 0.45 | 97.55 | 0 | 0 | 1.05 | 0.95 |
| IT | 1.01 | 0 | 96.15 | 0 | 2.84 | 0 |
| DT | 0 | 0 | 0 | 95.29 | 0 | 4.71 |
| US | 0 | 0 | 5.07 | 0 | 94.93 | 0 |
| DS | 1 | 0 | 0 | 4.01 | 0 | 94.99 |

Table 11. Normal shift (1.5–2.8) sigma with six features

| Pattern | NOR | CYC | IT | DT | US | DS |
|---|---|---|---|---|---|---|
| NOR | 100 | 0 | 0 | 0 | 0 | 0 |
| CYC | 0.47 | 99.52 | 0 | 0 | 0 | 0 |
| IT | 0 | 0 | 98.76 | 0 | 1.23 | 0 |
| DT | 0 | 0 | 0 | 98.65 | 0 | 1.34 |
| US | 0 | 0 | 1.72 | 0 | 98.27 | 0 |
| DS | 0 | 0 | 0 | 2.19 | 0 | 97.80 |

Table 12. Small shift less than (1.5) sigma with six features

| Pattern | NOR | CYC | IT | DT | US | DS |
|---|---|---|---|---|---|---|
| NOR | 99.9 | 0 | 0 | 0 | 0 | 0.05 |
| CYC | 0.45 | 99.55 | 0 | 0 | 0 | 0 |
| IT | 0 | 0 | 97.15 | 0 | 2.84 | 0 |
| DT | 0 | 0 | 0 | 97.29 | 0 | 2.70 |
| US | 0 | 0 | 3.06 | 0 | 96.93 | 0 |
| DS | 0 | 0 | 0 | 2.00 | 0 | 97.99 |

The above results can note that the network works better when decreasing the input dimensional from 13 to 6 features, which means the new approach (RCF-ANN) has significantly improved recognition accuracy. The classifier can detect small shifts with good recognition accuracy; all these results for developed patterns with an average of 10 runs.

### 3.4. Average run lengths

Average run lengths (ARL) are an important evaluation vector in SPC. (ARL0) calculates the stable process, which means how long the process is still stable before a false alarm, the large number of ARL0 is better than a small. (ARL1) for unstable process, it means how many observations are required before the correct unstable pattern is recognized. In this paper, the (ARL0) calculated for two different data sets for a stable process is equal (315) for normal shifting, and (260) for the small shifting dataset. ARL1 was calculated for two data sets (normal and small shifts). In normal shift, the (ARL1) equals (15), but in a small shift dataset, it is equal to (15.5). We can compare this work with previous work those used ANN-MLP as a classifier, as shown in Table 13.

We can note from Table 13 the recognition accuracy when using three feature selection algorithms; it gives good accuracy for diagnosing the normal pattern with 99.9% accuracy from the rest of the abnormal patterns within a small mean shifting data set (less than 1.5) sigma. Previous studies used the mean shifting data set (1.5–2.8) sigma. One of the possible reasons for higher recognition accuracy when compared to previous studies, even though they used ANN-MLP as a classifier, is that the six features selected for the first time and collected together as an input in this study improved recognition accuracy by presenting the raw data well. Using the mixed features (statistical and Shape) gives a good representation of data as an input to the network.

Table 13. Results comparison with previous studies

| Ref. | Model | Learning algorithm | Optimization | Input | Out. Patterns | AC % |
|---|---|---|---|---|---|---|
| Lu *et al*. [33] | MLP | Back-propagation | - | Statistical and shape features | 8 | 81.3 |
| Naeini and Bayati [6] | MLP | descending gradient | new feature belief variable | Statistical features | 6 | 97.36 |
| Hassan [31] | MLP | BFGS | Recognition only when necessary | Statistical features | 6 | 87.8 |
| This work | MLP | Back-propagation | - | Statistical and shape features | 6 | 98.88 |

## 4. CONCLUSION

The objective of this paper is to suggest an improved approach for feature selection by using three filter algorithms, namely, Relief, correlation, and Fisher. This provides more information than using one features selection algorithm to select the best features as input presentation. The ANN-BP was employed as a classifier to classify patterns from two datasets that had previously been utilised in research that were compared to this work. The first one with data set having a mean shifting between 1.5 σ to 2.8 σ, and the second data set with a small mean shifting (less than 1.5 σ). The results show RCF-ANN gave significantly better performance and good generalization. Also, we can note that the network has a good recognition accuracy for both the data sets and improved the correct recognition accuracy from 97.33% and 95.87% to 98.88% and 98.18% for normal and small shift, respectively. The mixed features (statistical and shape) significantly contributed and gave good accuracy. Note that some miss classification happened with small shift data because the small shift patterns became partly similar to the normal pattern. The average run length, ARL0 for the stable process is equal (315) for normal process, and (260) for the small shifting dataset, while the ARL1 improved when detecting a small shift to 15.5. This experimental result is limited to fully developed patterns. We plan to test the developing patterns with moving windows size and apply the different classifiers for future work.

**ACKNOWLEDGMENT**

The authors would like to thank the Ministry of Higher Education (MOHE), Universiti Teknologi Malaysia (UTM) for the financial support of this work. This research was funded Ministry of Higher Education (MOHE) through Universiti Teknologi Malaysia (UTM) under scheme UTM R&D Fund (4J506), UTM Shine (09G94) and UTM FR (22H06) and the Fundamental Research Grant Scheme (FRGS) [R.J130000.7851.5F316].

**REFERENCES**

[1] L. L. Boaventura, P. H. Ferreira, and R. L. Fiaccone, "On flexible statistical process control with artificial intelligence: classification control charts," *Expert Systems with Applications*, vol. 194, p. 116492, May 2022, doi: 10.1016/j.eswa.2021.116492.

[2] T. Kondakci and W. Zhou, "Recent applications of advanced control techniques in food industry," *Food and Bioprocess Technology*, vol. 10, no. 3, pp. 522–542, Mar. 2017, doi: 10.1007/s11947-016-1831-x.

[3] 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, Mar. 2020, doi: 10.1007/s10845-019-01473-0.

[4] O. El Farissi and H. Elboujaoui, "Improved multi-layer perceptron for recognition of control chart pattern," *International Journal of Computer Applications*, vol. 176, no. 36, pp. 39–42, Jul. 2020, doi: 10.5120/ijca2020920537.

[5] R. Haghighati and A. Hassan, "Recognition performance of imputed control chart patterns using exponentially weighted moving average," *European Journal of Industrial Engineering*, vol. 12, no. 5, pp. 637–660, 2018, doi: 10.1504/EJIE.2018.094599.

[6] M. K. Naeini and N. Bayati, "Pattern recognition in control chart using neural network based on a new statistical feature," *International Journal of Engineering*, vol. 30, no. 9, Sep. 2017, doi: 10.5829/idosi.ije.2017.30.09c.10.

[7] A. Addeh and B. M. Maghsoudi, "Control chart patterns detection using COA based trained MLP neural network and shape features," *Computational Research Progress in Applied Science & Engineering*, vol. 2, no. 1, pp. 5–8, 2016.

[8] M. Zaman and A. Hassan, "Improved statistical features-based control chart patterns recognition using ANFIS with fuzzy clustering," *Neural Computing and Applications*, vol. 31, no. 10, pp. 5935–5949, Oct. 2019, doi: 10.1007/s00521-018-3388-2.

[9] A. Hassan, M. S. N. Baksh, A. M. Shaharoun, and H. Jamaluddin, "Improved SPC chart pattern recognition using statistical features," *International Journal of Production Research*, vol. 41, no. 7, pp. 1587–1603, Jan. 2003, doi: 10.1080/0020754021000049844.

[10] R. Haghighati and A. Hassan, "Feature extraction in control chart patterns with missing data," *Journal of Physics: Conference Series*, vol. 1150, no. 1, p. 012013, Jan. 2019, doi: 10.1088/1742-6596/1150/1/012013.

[11] R. Russo, G. Romano, and P. Colombo, "Identification of various control chart patterns using support vector machine and wavelet analysis," *Annals of Electrical and Electronic Engineering*, vol. 2, no. 8, pp. 6–12, Aug. 2019, doi: 10.21833/AEEE.2019.08.002.

[12] W. Alwan, A. Hassan, and N. H. A. Ngadiman, "A review on input features for control chart patterns recognition," in *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 2021, pp. 2625–2635.

[13] X. Zhou, P. Jiang, and X. Wang, "Recognition of control chart patterns using fuzzy SVM with a hybrid kernel function," *Journal of Intelligent Manufacturing*, vol. 29, no. 1, pp. 51–67, Jan. 2018, doi: 10.1007/s10845-015-1089-6.

[14] M. Zhang, W. Cheng, and P. Guo, "Intelligent recognition of mixture control chart pattern based on quadratic feature extraction and SVM with AMPSO," *Journal of Coastal Research*, vol. 73, pp. 304–309, Mar. 2015, doi: 10.2112/SI73-053.1.

[15] M. Zhang, X. Zhang, H. Wang, G. Xiong, and W. Cheng, "Features fusion exaction and KELM with modified grey wolf optimizer for mixture control chart patterns Recognition," *IEEE Access*, vol. 8, pp. 42469–42480, 2020, doi: 10.1109/ACCESS.2020.2976795.

[16] M. Jamei *et al.*, "Combined terrestrial evapotranspiration index prediction using a hybrid artificial intelligence paradigm integrated with Relief algorithm-based feature selection," *Computers and Electronics in Agriculture*, vol. 193, 2022, doi: 10.1016/j.compag.2022.106687.

[17] M. R. Mahmood and M. B. Abdulrazzaq, "Performance evaluation of chi-square and Relief-F feature selection for facial expression recognition," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 27, no. 3, p. 1470, Sep. 2022, doi: 10.11591/ijeecs.v27.i3.pp1470-1478.

[18] H. Zhou, X. Wang, and R. Zhu, "Feature selection based on mutual information with correlation coefficient," *Applied Intelligence*, vol. 52, no. 5, pp. 5457–5474, Mar. 2022, doi: 10.1007/s10489-021-02524-x.

[19] Q. Gu, Z. Li, and J. Han, "Generalized Fisher score for feature selection," in *Proceedings of the 27th Conference on Uncertainty in Artificial Intelligence, UAI 2011*, Feb. 2011, pp. 266–273.

[20] D. Aksu, S. Üstebay, M. A. Aydin, and T. Atmaca, "Intrusion detection with comparative analysis of supervised learning techniques and Fisher score feature selection algorithm," in *Communications in Computer and Information Science*, vol. 935, 2018, pp. 141–149.

[21] R. F. Bartlett, "Linear modelling of pearson's product moment correlation coefficient: an application of Fisher's z-transformation," *The Statistician*, vol. 42, no. 1, p. 45, 1993, doi: 10.2307/2348110.

[22] M. Zaman and A. Hassan, "Fuzzy heuristics and decision tree for classification of statistical feature-based control chart patterns," *Symmetry*, vol. 13, no. 1, pp. 1–12, Jan. 2021, doi: 10.3390/sym13010110.

[23] 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, Aug. 2018, doi: 10.1016/j.isatra.2018.04.020.

[24] Z. Hong, Y. Li, and Z. Zeng, "Convolutional neural network for control chart patterns recognition," in *ACM International Conference Proceeding Series*, Oct. 2019, pp. 1–9, doi: 10.1145/3331453.3360974.

[25] X. Xu-Dong and M. Li-Qian, "Control chart recognition method based on transfer learning," in *Proceedings - 2018 4th Annual International Conference on Network and Information Systems for Computers, ICNISC 2018*, Apr. 2018, pp. 446–451, doi: 10.1109/ICNISC.2018.00097.

[26] 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, no. 3–2, pp. 5–9, 2017.

[27] A. Addeh, P. Zarbakhsh, S. J. S. Kharazi, and M. Harastani, "A hierarchical system for recognition of control chart patterns," in *Proceedings on 2018 International Conference on Advances in Computing and Communication Engineering, ICACCE 2018*, 2018, pp. 423–427, doi: 10.1109/ICACCE.2018.8441711.

[28] H. Malik, A. K. Yadav, F. P. G. Márquez, and J. M. Pinar-Pérez, "Novel application of Relief algorithm in cascaded artificial neural network to predict wind speed for wind power resource assessment in India," *Energy Strategy Reviews*, vol. 41, p. 100864, May 2022, doi: 10.1016/j.esr.2022.100864.

[29] C. Xue *et al.*, "Radiomics feature reliability assessed by intraclass correlation coefficient: a systematic review," *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 10, pp. 4431–4460, Oct. 2021, doi: 10.21037/qims-21-86.
[30] Y. Ramdhani, D. F. Apra, and D. P. Alamsyah, "Feature selection optimization based on genetic algorithm for support vector classification varieties of raisin," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 30, no. 1 pp. 192-199, Apr. 2023, doi: 10.11591/ijeecs.v30.i1.pp192-199.
[31] A. Hassan, "An improved scheme for online recognition of control chart patterns," *International Journal of Computer Aided Engineering and Technology*, vol. 3, no. 3–4, pp. 309–321, 2011, doi: 10.1504/IJCAET.2011.040050.
[32] J. Addeh, A. Ebrahimzadeh, M. Azarbad, and V. Ranaee, "Statistical process control using optimized neural networks: A case study," *ISA Transactions*, vol. 53, no. 5, pp. 1489–1499, Sep. 2014, doi: 10.1016/j.isatra.2013.07.018.
[33] Z. Lu, M. Wang, and W. Dai, "A condition monitoring approach for machining process based on control chart pattern recognition with dynamically-sized observation windows," *Computers and Industrial Engineering*, vol. 142, p. 106360, Apr. 2020, doi: 10.1016/j.cie.2020.106360.
[34] I. Kononenko, "Estimating attributes: analysis and extensions of RELIEF," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 784 LNCS, 1994, pp. 171–182, doi: 10.1007/3-540-57868-4_57.

## BIOGRAPHIES OF AUTHORS

**Waseem Alwan** received his B.S. degree in mechanical engineering and higher Diploma degree in Computer Aided in Instructional Engineering from University of Technology, Iraq, in 1999 and 2001, respectively. His M.S. degree in Mechanical and Manufacturing Engineering from Universiti Tun Hussein Onn Malaysia (UTHM), Malaysia, in 2019. He is currently a doctoral candidate in industrial engineering, Universiti Teknologi Malaysia. His research interests include control chart patterns recognition, features selection and missing data. He can be contacted at email: zaboon.w@graduate.utm.my.

**Dr. Nor Hasrul Akhmal Ngadiman** received his Doctor of Philosophy (Ph.D.) in Mechanical Engineering from Universiti Teknologi Malaysia (UTM) in 2016. He is currently Senior Lecturer in the Department of Materials, Manufacturing and Industrial Engineering, Faculty of Mechanical Engineering, UTM, Malaysia. He is a member of Institution of Mechanical Engineers and Chartered Mechanical Engineer. He is also Professional Engineer recognized by Board of Engineer Malaysia and Professional Technologist recognized by Malaysia Board of Technologist. Dr. Hasrul is involved (both as Project Leader and Project Member) in numerous research projects funded by the Ministry of Education and various industries as well as by the UTM. His papers have been published in both international and national journals. In addition to this. He can be contacted at email: norhasrul@utm.my.

**Asst. Prof. Dr. Adnan Hassan** is an Associate Professor in the School of Mechanical Engineering, Universiti Teknologi Malaysia (UTM). He received a B.Sc. (Hons.) in Industrial Engineering from the University of Miami, Florida, U.S.A (1986), a M.Sc. in Industrial Measurement Systems from Brunel University, U.K. (1992), and a Ph.D. in Mechanical Engineering from UTM (2003). From 2007 to 2009, he was the Head of Manufacturing and Industrial Engineering Department at UTM. From Nov 2009 till June 2011, he was seconded to King Abdul Aziz University (KAU), Saudi Arabia where he was the founding Chairman for the Department of Industrial Engineering, KAU Rabigh Branch. His research interest includes pattern recognition for process monitoring, supply chain, maintenance, and performance measure. He can be contacted at email: adnan@utm.my.

**Dr. Mohd Syahril Ramadhan Mohd Saufi** received Ph.D. degree in Faculty of Mechanical Engineering, faculty of Engineering from the Universiti Teknologi Malaysia, Malaysia. Now he works as senior lecturer at Faculty of Mechanical Engineering, faculty of engineering, Universiti Teknologi Malaysia, Skudai, Malaysia. His current research interests include vibration and acoustic emission analysis, machinery fault diagnosis and prognosis, and artificial intelligence. He can be contacted at email: mohdsyahrilramadhan@utm.my.

**Asst. Prof. Dr. Azanizawati Ma'aram** is an Associate Professor at the Faculty of Mechanical Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM). She obtained her Bachelor of Engineering (Mechanical-Industrial) and Master of Engineering (Advanced Manufacturing Technology) from Universiti Teknologi Malaysia, Malaysia. She pursued her Doctorate of Philosophy (Ph.D.) (Management) at the University of Liverpool, United Kingdom. She has held several positions including Director of Department of Materials, Manufacturing and Industrial Engineering, Associate Chair (Quality and Strategy), Head of Industrial Panel, Postgraduate Coordinator for Master of Science (Industrial Engineering), Laboratory Coordinator for Industrial Engineering, and as Director of Department. Her research interests include supply chain management, performance measurement, lean manufacturing, sustainability, ergonomics, safety and medical devices. She is currently active as a Project Leader and a Project Member on numerous research projects and has secured several grants funded by the university, the Ministry of Education (MoE) and industrial grants that involve hospitals, industry collaborators as well as international researchers. She can be contacted at email: niza@utm.my.

**Asst. Prof. Dr. Ibrahim Masood** is an Associate Professor in the Faculty of Mechanical and Manufacturing Engineering, Universiti Tun Hussein Onn Malaysia (UTHM). He received a B.Eng. (Hons.) in Mechanical Engineering (1997), a M.Eng. in Mechanical Engineering (Advanced Manufacturing Technology) (2004), and a Ph.D. in Mechanical Engineering from Universiti Teknologi Malaysia (UTM). From 1998 to 2001, he was appointed as Quality Control Engineer in a manufacturing industry. Since joined UTHM in August 2001 as an Instructor Engineer, he was promoted as a Lecturer in 2004, a Senior Lecturer in 2012 and an Associate Professor in 2018. Currently, he was the Head of Industrial Systems Engineering Focus Group at UTHM. His research interest includes quality engineering and control chart pattern recognition. He can be contacted at email: ibrahim@uthm.edu.my.