IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

SPECIAL SECTION ON ADVANCES IN MACHINE LEARNING AND COGNITIVE COMPUTING FOR INDUSTRY APPLICATIONS

Received August 12, 2019, accepted September 20, 2019, date of publication October 4, 2019, date of current version October 25, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2945596

A New Automatic Method for Control Chart Patterns Recognition Based on ConvNet and Harris Hawks Meta Heuristic Optimization Algorithm

NOORBAKHSH AMIRI GOLILARZ[1], ABDOLJALIL ADDEH[2], HUI GAO[1], LIAQAT ALI[3], AREF MORADKHANI ROSHANDEH[4], HAFIZ MUDASSIR MUNIR[5], AND RIAZ ULLAH KHAN[6]

[1]School of Computer Science and Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China
[2]Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol 47148-71167, Iran
[3]School of Information and Communication Engineering, University of Electronic Science and Technology of China, Chengdu, 611731, China
[4]Jacobs Engineering Group, Houston, TX 77072, USA
[5]School of Automation Engineering, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China
[6]Joint China-Cuba Laboratory For Frontier Research in Translational Neurotechnology, University of Electronic Science and Technology of China (UESTC), Chengdu 611731, China

Corresponding authors: Noorbakhsh Amiri Golilarz (noorbakhsh.amiri90@gmail.com) and Abdoljalil Addeh (jalil_addeh@yahoo.com)

This work was supported by the National Natural Science Foundation of China under Grant 61673085.

ABSTRACT The productions quality has become one of the essential issues in the modern manufacturing industry and several techniques have introduced for control and monitoring the production process. Control charts are the most practical and popular tools for continuously monitoring and, if required, make adjustments to the product or process. A new automatic method based on deep learning and optimization algorithms for nine control chart patterns (CCPs) recognition are proposed in this paper. This method has two principal parts: the classification part and the tuning part. In the last few years, a convolutional neural network (ConvNet) has led to an excellent performance on various tasks, like image processing, speech recognition, and signal processing. Therefore, in the classification part, ConvNet is used as the intelligent classifier for CCPs recognition. One significant difficulty of ConvNet is that it requires considerable proficiency to select suitable parameters like a number of kernels and their spatial sizes, learning rate, etc. The ConvNet parameters have domestic dependencies which make the tuning of these parameters a challenging task. According to these issues, in the tuning part of the proposed method, the Harris hawks optimization (HHO) algorithm is used for optimal tuning of ConvNet parameters. Contrasting the common CCPs recognition methods, the proposed method takes unprocessed data and passes to more than one hidden layer for extracting the optimal feature representation instead of relying on any feature engineering mechanisms. The quantitative and simulation results show the superiority of the proposed method over the previous techniques in terms of its performance.

INDEX TERMS Artificial intelligence, automation, neural network, optimization, pattern recognition.

I. INTRODUCTION

Every organization is trying to enhance its product quality at each step of the procedure of the manufacturing to obtain the global competitive advantage. One of the profitable tools of total quality control and management is statistical process control (SPC), which can be utilized for monitoring the process changes and also for enhancing the production quality [1], [2]. The role of SPC tools is illustrated in Fig. 1, which presents a process as a system with a set of inputs and an output. In the case of a manufacturing process, the controllable input factors $x_1, x_2, \ldots, x_p$ are process variable such as temperatures, pressures, feed rates, and other process variables. The inputs $z_1, z_2, \ldots, z_q$ are uncontrollable (or difficult to control) inputs, such as environmental factors or properties of raw materials provided by an external supplier. The production process transforms the input raw materials, component parts, and subassemblies into a finished

The associate editor coordinating the review of this manuscript and approving it for publication was Hong-Ning Dai.

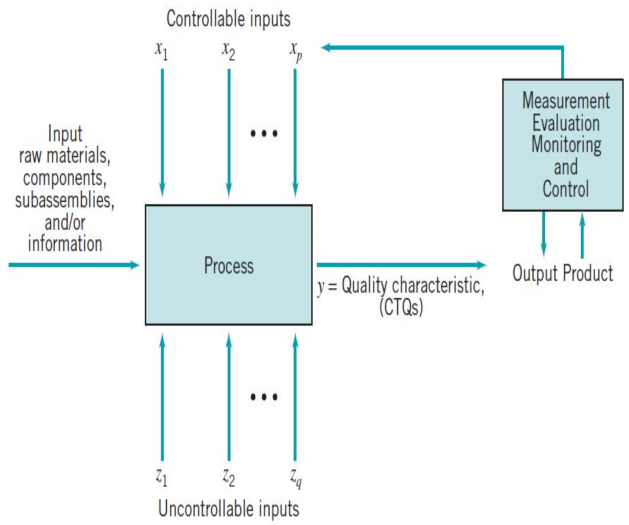

FIGURE 1. **Production process input and output [6].**

product that has several quality characteristics. The output variable y is a quality characteristic, that is, a measure of process and product quality. This model can also be used to represent nonmanufacturing or service processes [3]–[6].

Control charts are the most significant and popular tools in SPC to specify whether a process is running as planned or there are maybe some irregular causes of process variation [3]. Normal pattern (NOR) in the control charts indicates that the process is in its normal situation and there aren't any fault or variation. Except for the normal pattern, the reaming or abnormal patterns show that the process is out of control and the improvement actions are necessary. The irregular patterns are stratification (STA), cyclic (CYC), systematic (SYS), increasing trend (IT), decreasing trend (DT), the mixture (MIX), downward shift (DS) and upward shift (US) [6].

Accurate recognition of control chart patterns (CCP) are an important issue, because they monitor and control the production process. In recent years several methods have been proposed for accurate recognition and classification of CCPs including expert systems, support vector machine (SVM), different types of artificial neural networks (ANN) such as multilayer perceptron (MLP), radial basis function (RBF) and probabilistic neural networks (PNN), and fuzzy systems [7]–[42]. According to the previous researches for automatic CCPs recognition, there are three main aspects that should be considered during the design of recognizer. One of these issues is the feature extraction and feature selection. Based on the investigated papers, it is witnessed that using new features as the input of classifier have led to better performance [23]–[42].

Another issue is related to the type of classifier. In most studies, ANN, SVM, and ANFIS are used as the classifier. The most common and useful training algorithm for ANNs is the BP algorithm, that uses gradient information for finding the unknown weights and biases. Gradient-based algorithms easily get trapped in local minima notably for nonlinear and complicated pattern recognition problems. In addition, the convergence speed of gradient-based algorithms is very low even for simple pattern recognition problems. Also, there is not any standardized approach to choose the neural network's architecture. Generally, it is required to obtain this architecture empirically, which is a time-consuming process [43]. About the SVM, its accuracy is dependent on the selection of the kernel function and other parameters such as slack variables, cost parameter, and the margin of the hyperplane. There is a direct relationship between the failure in finding the optimal settings for an SVM model and its recognition accuracy. The computational cost of the SVM is another barrier [44], [45]. About the fuzzy systems, the error rate of fuzzy systems is high as they suffer from the necessity of large training data set and problems of random initial cluster center selection [46].

All the presented methods for CCPs recognition so far, demand the signal preprocessing, handcrafted features extraction and feature selection. The conventional techniques have acceptable performance, but they are very complicated and have several modules. In these systems, features are usually chosen using the trial-and-error or even by the experience. Therefore, we applied a convolutional neural network (ConvNet) in this paper to overwhelm the possible problems while using the traditional methods and also to acquire better detection accuracy without using any handcrafted feature extraction, feature selection features. In the proposed method, the Harris hawks optimization (HHO) algorithm is used for optimal tuning of ConvNet parameters.

The rest of this paper is established as follows: Section 2 represents ConvNet. Part 3 is about the optimization algorithm. In part 4, we presented the proposed method. Section 5, is the simulation results, and finally, part 6 is the conclusion.

II. ConvNet

The ConvNets are one of the newly introduced and useful machine learning tools which have excellent performance in different applications like pattern recognition and fault detection. A ConvNet consists of three sorts of main layers, containing pooling layer (Pool), convolutional layer (CONV), and fully-connected layer. The main structure of the ConvNet is shown in Figure 2. The ConvNet in Figure 2 has one CONV layer, one Pool layer, and one fully connected layer. The number of CONV and Pool layers or hidden layers can be more than one. The user should select the optimal number of hidden layers.

In CONV layer, some convolution kernels have been utilized to calculate new feature maps [47]. The value of each feature at the location (i, j) in the k-th feature map of the l-th layer, $z^{l}_{i,j,k}$ is given as follows:

$$z^{l}_{i,j,k} = W^{l^T}_{k} X^{l}_{i,j} + b^{l}_{k} \tag{1}$$

In Eq. (1), W^{l}_{k} represent the weight vector and b^{l}_{k} denotes the bias term of the k-th filter of the l-th layer, and

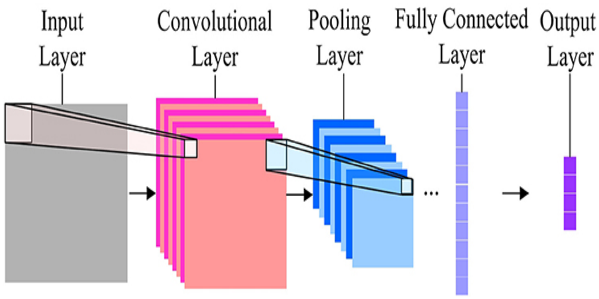

FIGURE 2. The main structure of the ConvNet.

$X_{i,j}^{l}$ shows the input patch centered at the position (*the* i,j) of the l-th layer. The activation value $a_{i,j,k}^{l}$ of convolutional feature $z_{i,j,k}^{l}$ is given by:

$$a_{i,j,k}^{l} = a\left(z_{i,j,k}^{l}\right) \tag{2}$$

where $a(.)$ is the activation function with the nonlinearity characteristic. Among several activation function types, the Rectified Linear Unit (ReLU) is one of the most effective and well-known activation function which can be defined as:

$$a_{i,j,k}^{l} = \max\left(z_{i,j,k}^{l}, 0\right) \tag{3}$$

A possible drawback of the ReLU activation function is that, whenever the unit is not active, it has zero gradients [48], [49]. Consequently, the training process will be slow because of the constant zero gradients. In [50] a new model of activation function called Exponential Linear Unit (ELU) is proposed to allow the faster learning procedure of ConvNet which results in better performance. Similar to ReLU, in ELU the vanishing gradient problem can be prevented properly by tuning the positive part of identity. Unlike the ReLU, there is a negative part in ELU to make the learning process fast. To lessen the units' variation while deactivated and make a robust ELU against noise, it is required to apply a saturation function as the negative part besides the unsaturated negative parts. The ELU is given by:

$$a_{i,j,k}^{l} = \max\left(z_{i,j,k}^{l}, 0\right) + \min\left(\lambda\left(e^{z_{i,j,k}} - 1\right), 0\right) \tag{4}$$

In this equation, λ is a free parameter that controls the saturation of ELU for negative inputs. Considering the advantages of the ELU activation function, the ELU has used in the proposed method for CCPs recognition.

Shift-invariance in the pooling layer can be acquired using diminishing the feature maps resolution. The Pooling layer lies between two CONV layers. The pooling function can be symbolized as $pool(.)$, for each feature map $a_{m,n,k}^{l}$ which is given as follows:

$$y_{i,j,k}^{l} = \text{pool}\left(a_{m,n,k}^{l}\right), \quad \forall (m,n) \in R_{ij} \tag{5}$$

In Eq. (5), R_{ij} is a local neighborhood around the location (i,j). The common pooling operations are L2-norm pooling, average pooling and, max pooling. After one or more hidden layer including CONV/Pool layers, there are fully connected and an output layer. The fully connected layer takes the extracted features by hidden layers and make a linear relationship between these new inputs and targets. Softmax is a popular used operator, especially for pattern recognition problems. Minimizing a suitable loss function which is defined on a particular task, makes it possible to get its optimum parameters (θ). If there exists N input-output relations $\{(x^{(n)}, y^{(n)}); n \in [1, 2, \ldots, N]\}$, where $x^{(n)}$ is the n-th input data, $y^{(n)}$ is its target label correspondingly and $o^{(n)}$ is the output of ConvNet. The loss of ConvNet can be stated as follows:

$$\ell = 1/N \sum_{n=1}^{N} \ell\left(\theta : y^{(n)}, o^{(n)}\right) \tag{6}$$

The training of ConvNet is a nonlinear and complicated optimization problem. To get the best fitting set of the parameters including the weights and biases, we should minimize the loss function. The common solution for optimizing the ConvNet network is Stochastic gradient descent [51].

III. HHO ALGORITHM

HHO is a new nature-inspired meta-heuristic optimizer that is motivated by Harris' hawks manner in nature for finding the food source. In this algorithm, several hawks all together attack a hunt to astonish it (exploration phase). Concerning the probability of evasion and running away for the hunt, the Hawks can perform several quick dives coming close the prey to surprise it and make it tired (exploitation phase). Based on the fleeing energy of the prey, the HHO algorithm can exchange its phase from exploration to exploitation and then, move between different manners of the exploitative phase. The hunt's energy can be diminished significantly through the running process. To model this behavior, the prey's energy can be defined as below:

$$E = 2E_0\left(1 - \frac{iter}{Iter_{\max}}\right) \tag{7}$$

where E denotes the running and fleeing energy of the hunt or prey at each iteration of the algorithm, $Iter_{\max}$ is the maximum number of iterations, and E_0 is the prey's initial energy. The value of $Iter_{\max}$ should be determined by the user. In this algorithm, E_0 randomly varies inside the interval $(-1, 1)$ at each iteration. When the E_0 value lessens from 0 to -1, it means that the hunt is exhausted and tired, with the E_0 value increases from 0 to 1, it indicates that the prey is reinforcing.

During the iterations, the hunt's energy for running away from Harris hawks tends to be diminished. When $|E| \geq 1$, the hawks search for different places to find the location of the prey, hence, the HHO carries out the exploration phase. On the other hand, when $|E| < 1$, the HHO carry out the exploitation phase. To model the process of attacking the target mathematically, based on its running manners and also the chasing patterns of the Harris' Hawks, four strategies including (1) Soft besiege, (2) Hard besiege, (3) Soft besiege with progressive rapid dives and (4) Hard besiege with progressive rapid dives are presented in the HHO [52]. The hawks will do a hard or soft besiege based on the prey's escaping behavior to trap it. In this case, the Hawks will surround and attack the prey from different directions and this process is dependent on the prey's retained power of escaping. After some iteration and tries, the fleeing target's energy will be lost gradually. Thus, it is the time that the besiege procedure can be intensified by the hawks so that they approach toward the tired prey and trap it or reach the final global answer. The pseudo-code of the HHO algorithm is shown in Figure 3. More details about the HHO algorithm can be found in [52].

IV. PROPOSED METHOD

In this study, a new hybrid method based on ConvNet is proposed for nine CCPs recognition. One of the benefits of the proposed method is feature representation that is learned automatically from the unprocessed training data, which is a significant difference from conventional hand-crafted feature representation methods. In the proposed technique, unprocessed data is fed to ConvNet, and the type of CCPs is recognized. The main drawbacks of applying ConvNet on a new task is the necessity of having the remarkable experience and ability to choose the appropriate and fitted parameters such as the type of activation function, learning rate, kernel sizes, number of kernels, etc. Because of having the internal dependencies of these parameters, they will be particularly expensive for tuning. For example, the higher number of kernels with small size leads to better performance; it also ruins the training due to a too large searching area. As another example, stride with optimal amount can reduce spatial resolution; leading to computational benefits and can reduce the overlap of receptive fields. Therefore the amount of these parameters should be appropriately selected.

To overcome this problem, in this paper the application HHO algorithm is introduced for the optimal design of ConvNet. In the proposed technique, each hawk shows the ConvNet parameters including:

1. Number of CONV/Pooling layers (NCPL)
2. Learning rate (α)
3. Kernels Number in the CONV layer (K_{CONV})
4. Kernels Size in the CONV layer (F_{CONV})
5. Stride in the CONV layer (S_{CONV})
6. Zero padding (P)
7. Type of Pooling Layer (TPL) including Max Pooling, Average Pooling or L2-norm Pooling.
8. Size of kernels in the Pooling layer (F_{POOL})
9. Stride in the CONV Pooling (S_{POOL})

In the proposed method, we determined the maximum number of CONV/Pooling layer equal to eight. The HHO should find the optimal number of CONV/Pooling layers in this range. In this method, the first array of each hawk represents the number of CONV/Pooling layer and the second array shows the learning rate. The reaming arrays represent the parameters of each layer. In each CONV/Pooling layer, there are seven parameters. Therefore if the first array of a sample hawk is 5, then we will have five CONV/Pooling layer, and the total number of unknown parameters is 37 $(2 + (5 \times 7) = 37)$. The first two parameters are a number of CONV/Pooling layers and learning rate. Or if the first array of a sample hawk is 8, then we will have eight CONV/Pooling layer, and the total number of unknown parameters is 58 $(2 + (8 \times 7) = 58)$.

V. SIMULATION RESULTS

In this part, the performance analysis of the proposed method has been evaluated. Here, several experiments have been done, and also we compared the results of the proposed method with other techniques available in the literature. The ConvNet is developed and evaluated in MATLAB (version 9.5[R2018b]) programming language on a PC with a Windows 10 64-bit professional and 64 gigabyte RAM. In this study, the proposed method and other classifiers are run five independent times, and the average of these five independent runs is reported as the final performance of each method.

A. DATASET PERFORMANCE ANALYSIS OF THE PROPOSED TECHNIQUE

To evaluate the performance analysis of the proposed method, 1000 samples for each pattern are produced using the equations presented in [4]. In these equations, r_i is the value of a standard normal variate at i-th $(i = 1, 2, \ldots, 60)$ time point and y_i is the observed value at i-th time point. Thus, various patterns of length 60 for a normal process with mean a (μ) and standard deviation (σ) can be produced by the following formulas:

1) Normal pattern:

$$y_i = \mu + r_i\sigma, \quad \mu = 80,\ \sigma = 5 \tag{8}$$

2) Stratification pattern:

$$y_i = \mu + r_i\sigma', \quad 0.2\sigma \leq \sigma' \leq 0.2\sigma \tag{9}$$

3) Systematic pattern:

$$y_i = \mu + r_i\sigma + d \times (-1)^i, \quad 1\sigma \leq d \leq 3\sigma \tag{10}$$

4) Mixture pattern:

$$y_i = \mu + r_i\sigma + (-1)^w \mathrm{m}, \quad 1.5\sigma \leq m \leq 2.5\sigma \,\&\, \text{w is 0 or 1} \tag{11}$$

where w is a binary integer value which is dependent to a random number $p(0 < \mathrm{p} < 1)$ and a pre-specified probability value $b = mp$, which determines the shifting between distributions. The value of b is fixed as 0.4, and thus, $w = 0 if\, \mathrm{p} < 0.4 \& \mathrm{w} = 1 if p \geq 0.4$

5) Cyclic pattern:

$$y_i = \mu + r_i\sigma + a\sin(2\pi i/T), \quad 1.5\sigma \leq a \leq 2.5\sigma \,\&\, 8 \leq T \leq 16 \tag{12}$$

In this equation, "a'' represents the amplitude of cyclic variation and "T'' shows the period of a cycle.

Inputs: The population size pop_{size} and the maximum number of iterations $Iter_{\max}$
Outputs: The position of the prey and corresponding fitness
Generate the initial population of hawks randomly $x_i = (i = 1, 2, \ldots, N)$
while (stopping criteria is not met) do
    Calculate the fitness values of hawks
    Set X_{rabbit} as the location of prey (best location)
    for (each hawk (Xi)) do
        Update the E_0 and jump strength J
        Update the E
        if (|E|≥ 1) then ▷Exploration
            Update the location vector
        if (|E|<1) then ▷Exploitation
            if (r≥0.5 and |E|≥ 0.5) then ▷Soft besiege
                Update the location vector
            else if (r≥0.5 and |E|<0.5) then ▷Hard besiege
                Update the location vector
            else if (r<0.5 and |E|≥ 0.5) then ▷Soft besiege with progressive rapid dives
                Update the location vector
            else if (r<0.5 and |E|<0.5) then ▷Hard besiege with progressive rapid dives
                Update the location vector
Return X_{rabbit}

FIGURE 3. Pseudo-code of HHO algorithm [52].

6) Increasing Trend pattern:

$$y_i = \mu + r_i\sigma + ig, \quad 0.05\sigma \leq g \leq 0.1\sigma \tag{13}$$

7) Decreasing Trend pattern:

$$y_i = \mu + r_i\sigma - ig, \quad -0.1\sigma \leq g \leq -0.05\sigma \tag{14}$$

In Eq. (14) and (15), "g" represents the magnitude of gradient for the trend patterns.

8) Upward Shift pattern:

$$y_i = \mu + r_i\sigma + ks, \quad k = 1 if\, \mathrm{i} \geq \mathrm{P},\ \mathrm{else\ k} = 0 \tag{15}$$

9) Downward Shift pattern:

$$y_i = \mu + r_i\sigma - ks, \quad k = 1 if\, \mathrm{i} \geq \mathrm{P},\ \mathrm{else\ k} = 0 \tag{16}$$

In Eq. (16) and (17), "k" is a parameter which determining the shift position and "s" represents the magnitude of the shift. We have generated 9000 samples $[Input]_{60\times 9000}$ (1000 sample per pattern). To show the performance analysis of the proposed system, 50% of the data is used for training and 50% for testing the proposed method.

B. PERFORMANCE ANALYSIS OF THE PROPOSED TECHNIQUE

In this subsection, the performance of the proposed technique is investigated. The parameterized functions can represent convnet architecture. To obtain the optimal values for parameters, we used the HHO algorithm. The population size (number of hawks) and the maximum number of iterations is set to 25 and 100, respectively. Based on the HHO algorithm, ConvNet with five hidden layers and learning rate equal to 0.00126 is chosen as the best ConvNet. The optimal values of the parameters of the best ConvNet are listed in Table 1. The input layer includes 60 neurons, and the hidden layer comprises five convolution and pooling layers with 4, 8, 12, 20 and 32 filters. The proposed method (HHO- ConvNet) with ELU activation functions can classify the CCPs with 99.80% accuracy. Tables 2 shows the confusion matrix of the proposed method. It is clear that the proposed method has high classification accuracy.

TABLE 1. The proposed architecture.

Layer	Layer type	Output shape	Kernel Size (F)	No. of Kernels (K)	Stride (S)	Zero padding (P)	No. Of trainable parameters
0		60×1	-	-	-	-	-
1	CONV	51×4	12×1	4	1	1	48
	POOL	46×4	6×1	4	1	-	-
2	CONV	40×8	9×1	8	1	2	288
	POOL	19×8	4×1	8	2	-	-
3	CONV	18×12	4×1	12	1	1	384
	POOL	16×12	3×1	12	1	-	-
4	CONV	9×20	4×1	20	2	2	960
	POOL	8×20	2×1	20	1	-	-
5	CONV	7×32	4×1	32	1	1	2560
	POOL	6×32	2×1	32	1	-	-

In this experiment, the performance of the ConvNet with the same architecture but with the ReLU activation function is estimated. Obtained results using ConvNet with ELU and ReLU activation function are listed in Table 3. The ConvNet with ReLU activation function has 99.55% accuracy. The standard deviation of the proposed method (HHO-ConvNet with ELU activation function) is zero, and the standard deviation of ConvNet with the ReLU activation function is ±0.14. It is obvious that choosing a suitable activation function is an important issue. The confusion matrix of this network is shown in Table 4.

TABLE 2. Confusion matrix of the proposed method with ELU activation function (99.80%).

	NOR	STR	SYS	MIX	CYC	IT	DT	US	DS
NOR	500	0	0	0	0	0	0	0	0
STR	0	500	0	0	0	0	0	0	0
SYS	0	0	499	0	1	0	0	0	0
MIX	2	0	0	498	0	0	0	0	0
CYC	0	0	0	0	500	0	0	0	0
IT	0	0	0	0	0	497	0	3	0
DT	0	0	0	0	0	0	499	0	1
US	0	0	0	0	0	0	0	500	0
DS	0	0	0	0	0	0	2	0	498

TABLE 3. Investigation of the activation function effect on ConvNet performance.

Classifier	Activation function	Accuracy (%)			Standard Deviation
		Min	Max	Mean	
ConvNet with optimal architecture	ELU	99.80	99.80	99.80	±0.0
ConvNet with optimal architecture	ReLU	99.11	99.75	99.55	±0.14

TABLE 4. Confusion matrix of the ConvNet with ReLU activation function (99.55%).

	NOR	STR	SYS	MIX	CYC	IT	DT	US	DS
NOR	499	0	0	1	0	0	0	0	0
STR	0	500	0	0	0	0	0	0	0
SYS	0	0	497	0	3	0	0	0	0
MIX	2	0	0	498	0	0	0	0	0
CYC	0	0	1	0	499	0	0	0	0
IT	0	0	0	0	0	498	0	2	0
DT	0	0	0	0	0	0	496	0	4
US	0	0	0	0	0	3	0	497	0
DS	0	0	0	0	0	0	4	0	496

The convergence of the HHO algorithm in finding the optimal architecture is shown in Figure 3. In this experiment, the HHO-ConvNet is run five independent times, and the convergence diagrams are plotted. In these five independent runs, the initial population is fixed. Therefore, all the experiments start from the same point. It is clearly proof enough that the proposed method has excellent convergence speed. As indicated in these figures, ConvNet accuracy gradually changed from iteration 0 to 100 and demonstrated no remarkable change after iteration 50. In fact, after 50 iterations, the algorithm has reached the best point, and the optimal parameters of ConvNet are found.

C. COMPARISON WITH OTHER CLASSIFIERS

To compare the performance analysis of the proposed technique with other machine learning algorithms, we have done several experiments. For this purpose, probabilistic neural networks (PNN), MLPNN with different training algorithms such as BP, Resilient propagation (Rprop) and Levenberg Marquardt (LM), RBFNN, ANFIS, SVM and random forest (RF) are considered. In these experiments, the unprocessed data is used as the input of different classifiers and the same ratio of training/testing is used.

MLPNN has some variables containing the number of neurons in each hidden layer, number of hidden layers, learning rate and type of transfer function which must be determined carefully. In RBFNN, the number of radial basis functions and their spreads have a high effect on the network's performance.

TABLE 5. Comparing the performance analysis of the proposed classifier with other classifiers using raw data.

Classifier	Input size	RA (%)	SD
PNN	60	94.71	±2.04
MLP (BP)	60	92.07	±3.28
MLP (RP)	60	95.73	±1.17
MLP (LM)	60	96.54	±1.07
RBFNN	60	96.87	±1.04
ANFIS	60	97.52	±0.46
SVM	60	97.32	±0.27
RF	60	96.49	±0.58
Proposed method	60	99.80	±0.0

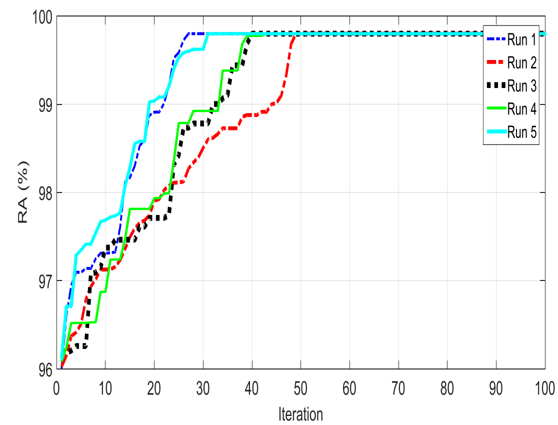

FIGURE 4. Convergence of the proposed method in different runs.

In PNN, the value of the spread is an important parameter that should be determined carefully. In ANFIS, the value of radii, type of membership function and fuzzy inference system must be determined accurately. In SVM, the type of kernel function, kernel function parameters, and penalty parameters must be determined carefully. In RF, the number of trees and the number of leaves have a high effect on its accuracy and performance. Therefore, in this experiment, the HHO algorithm is used to find the optimal value of these parameters. The obtained results are listed in Table 6. It is obvious that the proposed technique (HHO-ConvNet) has much better performance in comparison with other classifiers.

D. COMPARISON WITH DIFFERENT TECHNIQUES AVAILABLE IN THE LITERATURE

Control charts is one of the most practical and effective ways for monitoring the production process in many factories. Therefore this approach has been implemented in many industrial applications and there are vast studies in this area. In different studies, authors have used different dataset to investigate their method's performance. Therefore, there are not a unique dataset to compare the performance of different methods in similar conditions. Also, in different studies, different patterns and different samples are considered. Accordingly, in this section we have just mentioned the reported results in term of RA. Table 7 shows these comparison.

TABLE 6. Comparison with other methods.

Ref	Year	No. patterns	Input type	Classifier	Accuracy (%)
[23]	1997	Six	Feature (Shape)	MLPNN	99.00
[16]	2000	Six	Raw data	MLPNN	97.73
[28]	2003	Six	Features (Statistical)	MLPNN	96.79
[18]	2006	Six	Raw data	Optimized -LVQ	95.47
[20]	2006	Six	Raw data	WNN	97.70
[24]	2006	Eight	Feature (Shape)	MLPNN (LM)	96.13
[26]	2007	Eight	Feature (Shape)	CART	97.54
[25]	2009	Eight	Feature (Shape)	MLPNN (BP)	96.66
[29]	2010	Six	Feature (Shape + Statistical)	Optimized -SVM	99.58
[30]	2011	Six	Feature (Shape + Statistical)	MLPNN	99.21
[37]	2011	Six	Feature (Wavelet)	Optimized -SVM	99.37
[40]	2011	Six	Feature (geometric)	RBFNN	99.26
[4]	2012	Nine	Feature (Shape)	Expert system	95.21
[19]	2012	Eight	Raw data	SNN	98.61
[41]	2012	Six	Feature (geometric)	MLPNN	99.65
[31]	2013	Six	Feature (Shape + Statistical)	MLPNN-SVM	99.51
[32]	2013	Six	Feature (Shape + Statistical)	ANFIS	99.51
[38]	2013	Six	Feature (Wavelet)	ANFIS	99.32
[33]	2014	Six	Feature (Shape + Statistical)	RBFNN	99.58
[53]	2015	Nine	Features (Statistical)	SARG	89.90
[27]	2016	Six	Feature (Shape)	Optimized-MLPNN	99.21
[42]	2016	Six	Feature (fuzzy)	Optimized -SVM	99.68
[54]	2017	Seven	Feature (Shape + Statistical)	SVM	99.42
[35]	2018	Eight	Feature (Shape + Statistical)	Optimized -RBFNN	99.63
[34]	2018	Eight	Feature (Shape + Statistical)	MLPNN (RP)	99.50
This study	-	Nine	Raw data	HHO-ConvNet	99.80

In the control chart concept, pattern including normal, cyclic, trends and shift patterns are more common among the researchers and in many studies just these types of patterns have been studied. As a result, the proposed method for classification of six patterns are not applicable for classification eight or nine patterns. For instance, the proposed shape features in Ref [23] cannot be used for recognition of the stratification, systematic or mixture pattern form other patterns. All the feature-based introduced approaches can only differentiate the same patterns, and if we add a new pattern, the introduced method will not be practical [23]–[42], [53] and [54]. On the other hand, the approaches which have used unprocessed data as the input of classifier, have weak performance [4], [16], [18], and [19], with many real-life situations require a CCP recognition system to be proficient enough to discern all of the nine basic CCPs, including NOR, STR, SYS, MIX, IT, DT, US, and DS.

Furthermore, their CCP recognition system is not genuinely automated since some of the features extractions demand on the designer's inputs. For example, the limit points for the patterns segmentation into two windows need to be selected by the user. Based on the results obtained from Table 7, it is clear that only the proposed method (HHO-ConvNet) can recognize the nine CCPs with high accuracy.

VI. CONCLUSION

Considering the competitive market, controlling and monitoring of product quality has become one of the most critical issues in the manufacturing industry. In this paper a new automatic method based on ConvNet and optimization algorithm proposed for CCPs recognition. Using the proposed method, all the nine most commonly observed CCPs can be recognized automatically without any hand-crafted features. Also, the proposed method can be used for any arbitrary number of CCPs without any change. To see the quality and performance analysis of the proposed method (HHO-ConvNet), several experiments and comparisons with other methods performed. In these experiments, the optimal architecture and parameters of ConvNet are selected using HHO. The obtained results showed that the proposed technique has high accuracy and it can also classify the nine CCPs with the accuracy of 99.80. The results indicate the superiority of the proposed method over other classifiers such as MLPNN, RBFNN, ANFIS, RF, and SVM. Furthermore, the performance of the proposed method in terms of classification accuracy is much better than other reported results in the literature. According to the obtained results, this study highly recommends the application of HHO-ConvNet for CCPs recognition. Also, the proposed method can be used for other complicated pattern recognition problems such as heartbeats classification, breast cancer tumor type identification, etc.

REFERENCES

[1] J. Kim, G. M. Abdella, S. Kim, K. N. Al-Khalifa, and A. M. Hamouda, ''Control charts for variability monitoring in high-dimensional processes,'' *Comput. Ind. Eng.*, vol. 130, no. 2, pp. 309–316, 2019.

[2] F. A. P. Peres and F. S. Fogliatto, ''Variable selection methods in multivariate statistical process control: A systematic literature review,'' *Comput. Ind. Eng.*, vol. 115, no. 1, pp. 603–619, Jan. 2018.

[3] H. Hadian and A. Rahimifard, ''Multivariate statistical control chart and process capability indices for simultaneous monitoring of project duration and cost,'' *Comput. Ind. Eng.*, vol. 130, no. 2, pp. 788–797, Apr. 2019.

[4] M. Bag, S. K. Gauri, and S. Chakraborty, ''An expert system for control chart pattern recognition,'' *Int. J. Adv. Manuf. Technol.*, vol. 62, nos. 1–4, pp. 291–301, 2012.

[5] R.-S. Guh, ''Real-time pattern recognition in statistical process control: A hybrid neural network/decision tree-based approach,'' *Proc. Inst. Mech. Eng., B, J. Eng. Manuf.*, vol. 219, no. 3, pp. 283–298, 2005.

[6] D. C. Montgomery, *Introduction to Statistical Quality Control*. Hoboken, NJ, USA: Wiley, 2009.

[7] A. J. Duncan, *Quality Control and Industrial Statistics*. Chicago, IL, USA: Irwin, 1986.
[8] E. L. Grant and R. S. Leavenworth, *Statistical Quality Control*. New York, NY, USA: McGraw-Hill, 1996.
[9] L. S. Nelson, "The Shewhart control chart—Tests for special causes," *J. Qual. Technol.*, vol. 16, no. 3, pp. 237–239 1984.
[10] Western Electric Company Staff, *Statistical Quality Control Handbook*. Indianapolis, IN, USA: Mack Printing Company, 1958.
[11] C.-S. Cheng and N.-F. Hubele, "Design of a knowledge-based expert system for statistical process control," *Comput. Ind. Eng.*, vol. 22, no. 4, pp. 501–517, 1992.
[12] T. Lee Lucy-Bouler, "Using autocorrelations, CUSUMs and runs rules for control chart pattern recognition: An expert system approach," Ph.D. dissertation, Dept. Tuscaloosa, Alabama Univ., Tuscaloosa, AL, USA, 1991.
[13] J. A. Swift and J. H. Mize, "Out-of-control pattern recognition and analysis for quality control charts using Lisp-based system," *Comput. Ind. Eng.*, vol. 28, no. 1, pp. 81–91, 1995.
[14] R.-S. Guh, "A hybrid learning-based model for on-line detection and analysis of control chart patterns," *Comput. Ind. Eng.*, vol. 49, pp. 35–62, Aug. 2005.
[15] S. Du, D. Huang, and J. Lv, "Recognition of concurrent control chart patterns using wavelet transform decomposition and multiclass support vector machines," *Comput. Ind. Eng.*, vol. 66, no. 2, pp. 683–695, 2013.
[16] S. Sağiroğlu, E. Beşdok, and M. Erler, "Control chart pattern recognition using artificial neural networks," *Turkish J. Elect. Eng.*, vol. 8, no. 1, pp. 137–148, 2000.
[17] M. B. Perry, J. K. Spoerre, and T. Velasco, "Control chart pattern recognition using back propagation artificial neural networks," *Int J. Prod. Res.*, vol. 39, no. 15, pp. 3399–3418, 2011.
[18] D. T. Pham, S. Otri, A. Ghanbarzadeh, and E. Koc, "Application of the bees algorithm to the training of learning vector quantisation networks for control chart pattern recognition," in *Proc. 2nd Int. Conf. Inf. Commun. Technol.*, Damascus, Syria, 2006, pp. 1624–1629.
[19] M. H. A. Awadalla and M. A. Sadek, "Spiking neural network-based control chart pattern recognition," *Alexandria Eng. J.*, vol. 51, no. 2, pp. 27–35, 2012.
[20] S. Wu and B. Wu, "Wavelet neural network-based control chart patterns recognition," in *Proc. 6th World Congr. Intell. Control Automat.*, Dalian, China, Jun. 2006, pp. 9718–9721.
[21] H. Tang, H. Liu, W. Xiao, and N. Sebe, "Fast and robust dynamic hand gesture recognition via key frames extraction and feature fusion," *Neurocomputing*, vol. 331, no. 2, pp. 424–433, 2019.
[22] W. Zhong, L. Jiang, T. Zhang, J. Ji, and H. Xiong, "Combining multilevel feature extraction and multi-loss learning for person re-identification," *Neurocomputing*, vol. 334, no. 2, pp. 68–78, 2019.
[23] D. T. Pham and M. A. Wani, "Feature-based control pattern recognition," *Int. J. Prod. Res.*, vol. 35, no. 3, pp. 1875–1890, 1997.
[24] S. K. Gauri and S. Chakraborty, "Feature-based recognition of control chart patterns," *Comput. Ind. Eng.*, vol. 51, no. 2, pp. 726–742, 2006.
[25] S. K. Gauri and S. Chakraborty, "Recognition of control chart patterns using improved selection of features," *Comput. Ind. Eng.*, vol. 56, no. 4, pp. 1577–1588, 2009.
[26] S. K. Gauri and S. Chakraborty, "A study on the various features for effective control chart pattern recognition," *Int. J. Adv. Manuf. Technol.*, vol. 34, nos. 3–4, pp. 385–398, 2007.
[27] A. Addeh and B. M. Maghsoudi, "Control chart patterns detection using COA based trained MLP neural network and shape features," *Comput. Res. Prog. Appl. Sci. Eng.*, vol. 2, no. 1, pp. 5–8, 2016.
[28] A. Hassan, M. S. N. Baksh, A. M. Shaharoun, and H. Jamaluddin, "Improved SPC chart pattern recognition using statistical features," *Int. J. Prod. Res.*, vol. 41, no. 7, pp. 1587–1603, 2003.
[29] V. Ranaee, A. Ebrahimzadeh, and R. Ghaderi, "Application of the PSO–SVM model for recognition of control chart patterns," *ISA Trans.*, vol. 49, no. 3, pp. 577–586, 2010.
[30] J. Addeh, A. Ebrahimzadeh, and V. Ranaee, "Control chart pattern recognition using adaptive back-propagation artificial neural networks and efficient features," in *Proc. 2nd Int. Conf. Control, Instrum. Automat. (ICCIA)*, Shiraz, Iran, 2011, pp. 742–746.
[31] A. Ebrahimzadeh, J. Addeh, and V. Ranaee, "Recognition of control chart patterns using an intelligent technique," *Appl. Soft Comput.*, vol. 13, no. 2, pp. 2970–2980, 2013.
[32] J. Addeh, A. Ebrahimzadeh, and H. Nazaryan, "A research about pattern recognition of control chart using optimized ANFIS and selected features," *J. Eng. Technol.*, vol. 3, no. 1, pp. 6–16, 2013.
[33] J. Addeh, A. Ebrahimzadeh, M. Azarbad, and V. Ranaee, "Statistical process control using optimized neural networks: A case study," *ISA Trans.*, vol. 53, no. 5, pp. 1489–1499, 2014.
[34] A. Addeh, P. Zarbakhsh, S. J. S. Kharazi, and M. Harastani, "A hierarchical system for recognition of control chart* patterns," in *Proc. Int. Conf. Adv. Comput. Commun. Eng. (ICACCE)*, Paris, France, 2018, pp. 423–427.
[35] A. Addeh, A. Khormali, and N. A. Golilarz, "Control chart pattern recognition using RBF neural network with new training algorithm and practical features," *ISA Trans.*, vol. 79, no. 2, pp. 202–216, Aug. 2018.
[36] K. Assaleh and Y. Al-Assaf, "Features extraction and analysis for classifying causable patterns in control charts," *Comput. Ind. Eng.*, vol. 49, no. 2, pp. 168–181, 2005.
[37] V. Ranaee and A. Ebrahimzadeh, "Control chart pattern recognition using a novel hybrid intelligent method," *Appl. Soft Comput.*, vol. 11, no. 2, pp. 2676–2686, 2011.
[38] A. B. Bayat, A. Gharekhani, M. A. Mohajeran, and J. Addeh, "Control chart patterns recognition using optimized adaptive neuro-fuzzy inference system and wavelet analysis," *J. Eng. Technol.*, vol. 3, no. 2, pp. 76–81, 2013.
[39] C.-H. Wang and W. Kuo, "Identification of control chart patterns using wavelet filtering and robust fuzzy clustering," *J. Intell. Manuf.*, vol. 18, no. 2, pp. 343–350, 2007.
[40] J. Addeh, A. Ebrahimzadeh, and V. Ranaee, "Application of the PSO-RBFNN model for recognition of control chart patterns," in *Proc. 2nd Int. Conf. Control, Instrum. Autom. (ICCIA)*, Shiraz, Iran, 2011, pp. 747–752.
[41] A. Ebrahimzadeh, J. Addeh, and Z. Rahmani, "Control chart pattern recognition using K-MICA clustering and neural networks," *ISA Trans.*, vol. 51, no. 1, pp. 111–119, Jan. 2012.
[42] A. Khormali and J. Addeh, "A novel approach for recognition of control chart patterns: Type-2 fuzzy clustering optimized support vector machine," *ISA Trans.*, vol. 63, no. 2, pp. 256–264, Jul. 2016.
[43] J. R. Zhang, J. Zhang, T.-M. Lok, and M. R. Lyu, "A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training," *Appl. Math. Comput.*, vol. 185, no. 1, pp. 1026–1037, 2007.
[44] C. Campbell and N. Cristianini, "Simple learning algorithms for training support vector machines," Dept. Eng. Math., Univ. Bristol, Bristol, U.K., Tech. Rep., 1998.
[45] C. Zhang, P. Li, A. Rajendran, Y. Deng, and D. Chen, "Parallelization of multicategory support vector machines (PMC-SVM) for classifying microarray data," *BMC Bioinf.*, vol. 7, no. 2, p. S15, Des. 2006.
[46] D. J. Hemanth, C. K. S. Vijila, and J. Anitha, "Application of neuro-fuzzy model for mr brain tumor image classification," *Int. J. Biomed. Imag.*, vol. 16, no. 1, pp. 95–102, 2009.
[47] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR).*, Jun. 2015, pp. 1–9.
[48] A. L. Maas, A. Y. Hannun, and A. Y. Ng, "Rectifier nonlinearities improve neural network acoustic models," in *Proc. Int. Conf. Mach. Learn. (ICML).*, vol. 30, 2013, p. 3.
[49] M. D. Zeiler, M. Ranzato, R. Monga, M. Mao, K. Yang, Q. V. Le, P. Nguyen, A. Senior, V. Vanhoucke, and J. Dean, "On rectified linear units for speech processing," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP).*, 2013, pp. 3517–3521.
[50] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUs)," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2016, pp. 1–14.
[51] M. Zinkevich, M. Weimer, L. Li, and A. J. Smola, "Parallelized stochastic gradient descent," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS).*, 2010, pp. 2595–2603.
[52] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," *Future Gener. Comput. Syst.*, vol. 97, no. 3, pp. 849–872, Aug. 2019.
[53] K. Lavangnananda and S. Khamchai, "Capability of control chart patterns classifiers on various noise levels," *Procedia Comput. Sci.*, vol. 69, no. 2, pp. 26–35, 2015.
[54] 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 Eng.*, vol. 174, no. 3, pp. 281–288, 2017.