Refrigerated Showcase Fault Detection by a Pasting based Artificial Neural Networks using Parallel Multi-population Modified Brain Storm Optimization and Correntropy

Naoya Otaka and Yoshikazu Fukuyama
Graduate School of Advances Mathematical Sciences
Meiji University
Tokyo, Japan
yfukuyam@meiji.ac.jp

Yu Kawamura, Kenya Murakami, Adamo Santana,
Tatsuya Iizaka, and Tetsuro Matsui
Corporate R&D Headquarters
Fuji Electric Co., Ltd.
Tokyo, Japan

Abstract— **This paper proposes a fault detection method for refrigerated showcase by Pasting based Artificial Neural Networks (PANNs) using Parallel Multi-population Modified Brain Storm Optimization (PMP-MBSO) and Correntropy. PMP-MBSO is an evolutionary computation technique and Correntropy is one of loss functions (LFs) for ANN. At present, in Japan, the number of convenience stores is about 50,000, and refrigerated showcases placed in each store have various characteristics. Since it is difficult to adjust fault detection methods for all different refrigerated showcases, an automatic parameter adjustment method for refrigerated showcase systems is required. The proposed PANNs using PMP-MBSO and Correntropy can realize the automatic adjustment. Using actual operation data of showcases, detection accuracy of the proposed method is verified to be higher than those of the ANNs using least square error (LSE) and stochastic gradient descent (SGD), and the ANNs using Correntropy and MBSO. Effectiveness of the proposed method is verified by applying Friedman test. It is verified to speed up the proposed method as well.**

Keywords— fault Detection, refrigerated showcase, artificial neural network, correntropy, parallel multi-population modified brain storm optimization, pasting

I. INTRODUCTION

In refrigerated showcases placed in convenience stores and supermarkets, fans send cool air down to the bottom of the showcases in order to create air curtains. The air curtains prevent intrusions of outside air, and keep perishable products and beverages at constant temperatures. However, in rare cases, temperatures inside refrigerated showcases may not be maintained properly due to faults of the showcases such as frost formations or refrigerant leakage. This may cause negative impacts on customers. For this reason, accurate fault detection as early as possible is essential in order to maintain customer service.

In an early stage of researches, to the best of authors' knowledge, there are no researches especially on fault detection for refrigerated showcases. Since same refrigeration cycles are utilized in refrigerated showcases and air conditioners, the similar fault detection methods must be developed for both systems. As fault detection methods for air conditioners, various methods using classical artificial intelligence and an integrated method of multiple techniques have been proposed. Classical artificial intelligence have utilized in rule-based approaches [1-3]. As an integration method, a method combining principal component analysis and support vector machine has been proposed [4]. At present, in Japan, the number of convenience stores is about 50,000, and refrigerated showcases installed in each store have various characteristics [5]. Classical artificial intelligence and the integration method require to adjust rules and parameters of the methods for each store. Therefore, it is difficult for experts to adjust the methods for all showcase systems with different characteristics in practical situations.

Machine learning techniques are able to create rules and models with only data automatically without using expert knowledge and understanding of specialized techniques. Therefore, machine learning based methods have been proposed as fault detection methods for refrigerated showcases. A. Santana proposed fault detection methods using unsupervised learning and supervised learning [6-7]. In the papers, comparing various techniques, effectiveness of one of the machine learning techniques, namely ANN based methods for fault detections of refrigerated showcases has been confirmed [6-7]. However, these methods have a problem that these ANN are trained using only correctly measured data. Namely, it assumed that measurement systems are correctly operated.

In practical situations, incorrectly measured data may exist in measurement data from various sensors including missing values due to various factors. LSE utilized as a LF of ANN is known to be affected by such data [8]. Therefore, Using LSE, there is a possibility that accuracy of fault detection for test data may be deteriorated. Therefore, in practical situations, it is necessary for engineers to remove incorrectly measured data from training data in advance. It is time-consuming for engineers and causes operation cost rise of fault detection systems. Correntropy has been confirmed to solve this challenge [8-13]. When Correntropy is utilized as a LF, training is not affected by incorrectly measured data. In addition, parameter adjustment of ANN has been realized by SGD conventionally. However, due to optimization only using a single point, SGD may not be able to escape from local minima practically. It has been confirmed that evolutionary computation techniques based on multipoint search can escape from local minima for parameter adjustment of ANN [14]. MBSO is one of new evolutionary computation techniques improving BSO algorithm [15]. Effectiveness of MBSO has been confirmed for application of various fields [16-18]. Considering above points of view, authors have been confirmed that application of Correntropy and MBSO enables high-precision training that is not affected by incorrectly measured data [10]. However, there is a possibility that accuracy of fault detection may be improved. In addition, in a practical situation, fault detection of refrigerated showcases is realized by a cloud service in data centers. The service has to treat huge numbers of refrigerated showcases in stores and huge number of ANNs have to be trained. Therefore, it is necessary to speed up training time in practical systems.

Pasting is one of the methods for improving accuracy of fault detection [19]. Effectiveness of Pasting has been verified in various problems using machine learning [20, 21]. Moreover, Parallel Multi-population (PMP) is one of the methods for improving solution quality and computation time for evolutionary computation methods. Therefore, Applying PMP to MBSO, namely PMP-MBSO may be one of the way to improve accuracy of fault detection and speed up training time of ANNs. PMP-MBSO has been proposed in [22] and effectiveness of PMP-MBSO has been verified in a practical optimization problem [22].

This paper proposes a fault detection method for refrigerated showcases by PANNs using PMP-MBSO and Correntropy. Contributions of this paper can be summarized as follows:

- Proposal of a novel fault detection method for fault detection of refrigerated showcases using Pasting based ANNs using PMP-MBSO and correntropy.
- Verification of effectiveness of the proposed method by comparing with the conventional ANNs with SGD and LSE [7] and ANNs with MBSO and Correntropy [10] using actual operation data of refrigerated showcases.

II. REFRIGERATED SHOWCASE FAULT DETECTION

To maintain constant temperatures of refrigerated showcases, a large amount of refrigerated showcase data including a flow rate of refrigerant is measured in the refrigerated showcases. Showcase data can be applied to fault detection as ANN's training data. For example, as shown in Fig. 1, measured showcase data are collected in data center through Internet. Therefore, in the data center, various showcase fault detection models can be created offline. Showcase states can be detected online using a fault detection model, and the model can be used to detect online data. If the system is implemented, abnormal status of showcases can be detected through Internet as soon as the abnormal status appear. In such cases, service engineers need to go to the stores immediately and inspect the showcases. Even if normal status are incorrectly judged as abnormal status in particular store's showcases, service engineers have to go to the store and inspect the showcase every time. These are waste of actions and increase operating costs. If abnormal status are incorrectly judged as normal status in particular store's showcases, and if store clerks are unable to detect the abnormal status, perishables may be degraded and customers may lose their trust in the store.

III. CONVENTIONAL PARAMETER ADJUSTMENT METHODS OF ANN

Figure 2 shows a typical structure of a three-layered ANN. As shown in the figure, The ANN Outputs (o_{pk}) are calculated using forward propagation. In forward propagation, summations of input values multiplied by ANN parameters (w_{ij}, w_{jk}) are input to an activation function and an output value is calculated using the following equations [23]:

$$u_n = \sum_{m=0}^{M} w_{mn} a_m \tag{1}$$

$$b_n = \frac{1}{1 + e^{-u_n}} \tag{2}$$

where u_n is an internal value of the nth unit, M is the number of units at the target layer, w_{mn} is a parameter between the mth and nth units between two layers, a_m is an input value of the mth unit, b_n is an output value of the nth unit.

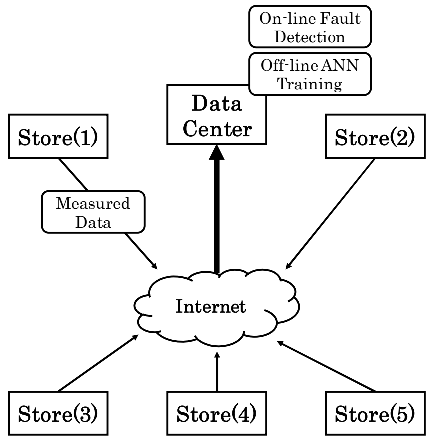


Fig.1 An example of a fault detection system of refrigerated showcases through Internet.

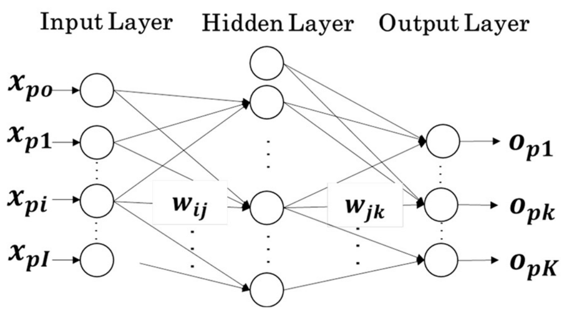


Fig.2 A typical structure of a three-layered ANN.

Conventionally, SGD with LSE has been applied to adjust ANN parameters (w_{ij}, w_{jk}) . A LF using LSE is shown below:

$$L_p = \frac{1}{2}\sum_{k=1}^{K}\left(t_{pk} - o_{pk}\right)^2 \quad (p = 1, \cdots, P) \tag{3}$$

where L_p is the pth LF, t_{pk} is the pth target data of the kth output unit, o_{pk} is the pth output of the kth output unit, P is the number of training data.

If LSE is utilized as the LF for ANN, loss function values (LFVs) using errors between outputs and target data of ANN are shown in Fig. 3. As shown in the figure, LFVs increase significantly when the error becomes large using LSE. The ANN using LSE solves a minimization problem. Usually, using correctly measured data, the ANN parameters are gradually adjusted to appropriate values. However, in practical situations, incorrectly measured data may exist in measurement data from various sensors including missing values due to various factors. When incorrectly measured data are utilized as training data, adjustment of parameters is greatly impacted from the incorrectly measured data. Consequently, an inappropriate model may be created for test data. In other words, a trained model may be incorrect because the model is adjusted to fit the incorrectly measured data. An impact on a fault detection problem with incorrectly measured data using LSE is shown below. Figure 4 shows two stages for the fault detection problem by an ANN using LSE. Using LSE, a decision boundary is affected by the incorrectly measured data as shown in Fig. 4(a). Thus, as shown in Fig. 4(b), test data are incorrectly classified using the bent decision boundary at a test stage. Treatment of incorrectly measure data is the first challenge for fault detection in practical situations. Correntropy has been confirmed to solve this challenge [8-13].

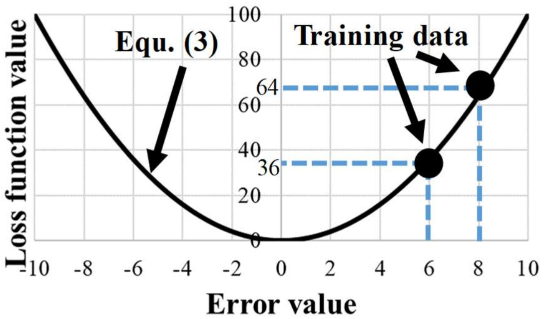


Fig.3. LFVs using errors between target values and outputs of ANN when LSE is applied to the LF.

Following equation is utilized for update of ANN parameters using SGD:

$$\Delta w_{mn}(t) = -\frac{\partial L_p(t)}{\partial w_{mn}(t)} \quad (t = 1, \cdots, T) \tag{4}$$

$$w_{mn}(t+1) = w_{mn} + \eta \Delta w_{mn} \tag{5}$$

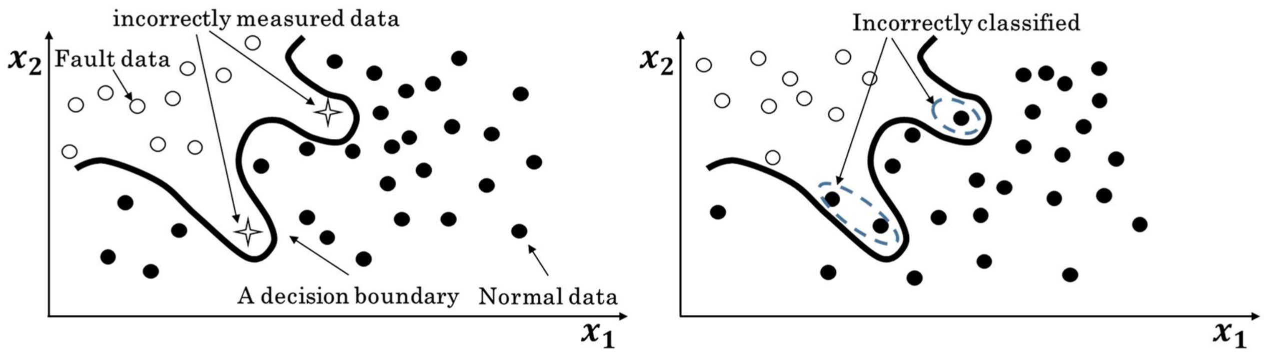
where $w_{mn}(t)$ is a parameter between the mth unit and nth unit at epoch t, T is the maximum number of epochs, $\Delta w_{mn}(t)$ is a gradient of $w_{mn}(t)$ at epoch t, η is a learning rate.

In parameter adjustment using SGD, Equ. (4) is calculated by a partial derivative using a different LF for each pattern in Equ. (3). Therefore, ANN parameters can avoid local minimum solutions. However, practically, ANN training may be trapped in local minimum solutions because SGD utilizes a single searching point. For this reason, researches using evolutionary computation techniques with multiple searching points have been conducted to improve detection accuracy [14]. In order to tackle the two challenges, the new method is proposed and shown in the next section.

IV. The Proposed Fault Detection Method for Refrigerated Showcase by a Correntropy Based PANNs using PMP-MBSO

A. *Overview of Correntropy for a loss function of ANN*

As observed in section 3, when errors between target values and outputs increase, LFVs using LSE increase significantly. Therefore, test evaluation is affected to fit incorrectly measured data. Consequently, ANN parameters have to be adjusted using only correctly measured training data. In order to tackle the problem, Correntropy has been developed by W. Liu, et al. in 2006 [24]. Even if incorrectly measured data exist in training data, ANN parameters are correctly adjusted using a LF with Correntropy regardless of the incorrectly measured data. Following equation is applied to a LF of ANN using Correntropy:



(a) A training stage using LSE. (b) A test stage using LSE.

Fig. 4. Inappropriate training at the stage and incorrect detection at the test stage using LSE.

$$maxL = \frac{1}{P \times K} \sum_{p=1}^{P} \sum_{k=1}^{K} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\left(t_{pk} - o_{pk}\right)^2}{2\sigma^2}\right) \quad (6)$$

where σ is a kernel size.

LFVs using Correntropy are maximized by errors between target values and outputs of ANN because normal distribution formula is utilized for Correntropy.

When Correntropy is applied to the LF, LFVs using errors between target values and outputs are shown in Fig. 5. The ANN using Correntropy solves a maximization problem. As observed in Fig. 5, LFVs become small significantly as the errors increase. Hence, training is not impacted by the huge errors even if incorrectly measured data exist in training data. Namely, training is not impacted by the incorrectly measured data even if incorrectly measured data exist in training data.

An impact on fault detection problems including incorrectly measured data by the ANN using Correntropy is shown below. Figure 6 shows test and training stages for a fault detection problem by the ANN using Correntropy. Using Correntropy, incorrectly measured data does not affect the decision boundary in Fig. 6(a). Namely, ANN can be properly trained. Thus, incorrect detection may be prevented by the uncomplex decision boundary in Fig. 6(b). For load forecasting of power systems, effectiveness of the Correntropy based ANN using training data including incorrectly measured data has been confirmed [13].

A kernel size (σ) in Equ. (6) is a hyper parameter that decides a tolerance of the errors. In other words, the impact of incorrectly measured data is highly dependent on the kernel size. Consequently, the kernel size should be decided according to a target issue.

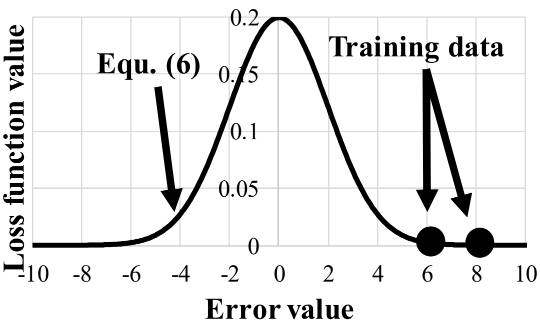


Fig.5. LFVs using errors between target values and outputs of ANN when Correntropy is applied to the LF.

B. Overview of Pasting

Pasting is an ensemble learning method using subsets created from original data and it has been proposed by L. Breiman in 1999 [19]. Typical ensemble learning methods using subsets created from original data are Bagging and Pasting. Bagging is a method using subsets created from original data with duplication, while Pasting is a method using subsets created from original data without duplication. An appropriate method for a certain problem depends on the problem. An algorithm of Pasting is shown below. As observed in Fig. 7, firstly, a plurality of independent subsets is created from original data without duplication. Then, classifiers are created for each subset using a same learning algorithm. In test stage, output values of each ANN are calculated using test data. Then, a final result is calculated with output values of each ANN using a certain ensemble rule such as decision by majority. Comparing various rules, this paper utilizes decision by majority as the ensemble rule.

C. Overview of PMP-MBSO

1) Overview of MBSO

MBSO is a new evolutionary computation technique improving BSO algorithm [15, 25]. Main processes utilized in MBSO have the following four steps.

- Step 1 Individuals are initially generated randomly within constraints.

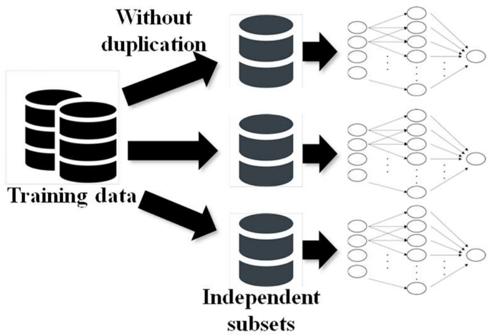
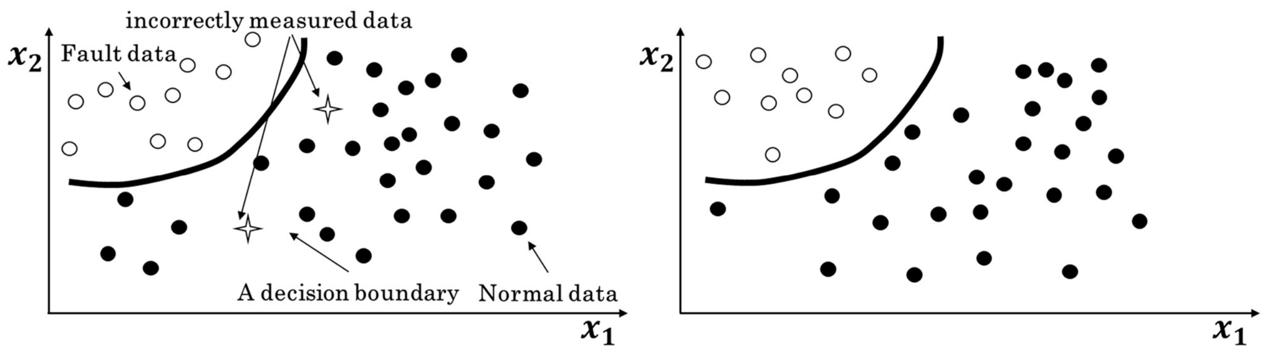


Fig.7 A training stage using PANN.



(a) A training stage using Correntropy. (b) A test stage using Correntropy.

Fig. 6. Appropriate training at the stage and correct detection at the test stage using Correntropy.

- Step 2 Several clusters are created by dividing whole individuals using simple grouping method (SGM). In a practical situation of fault detection of refrigerated showcases, a cloud service is performed in data centers. The service has to treat huge numbers of refrigerated showcases in stores and huge number of ANNs have to be trained. Therefore, saving computation time is really important in practical systems. In new individuals' generation, sharing each best individual's information is important. Consequently, new individuals' generation is not influenced from accuracy of clustering. Therefore, decline of solution quality may be prevent using SGM and MBSO can speed up computation time for practical applications.

- Step 3 Using the following equations, newly individuals are updated:

$$y_{ij}^{iter} = \begin{cases} rand(L_j, H_j) & if\ rand(0,1) < p_r \quad (7) \\ x_{ij}^{iter} + rand(0,1) \times (x_{aj}^{iter} - x_{bj}^{iter}) & otherwise \quad (8) \end{cases}$$

$$(i = 1, \cdots, N_I, j = 1, \cdots, N_{DV})$$

where y_{ij}^{iter} is a value of decision variable j of updated new individual i at iteration $iter$, L_j and H_j are the lower and upper bounds of decision variable j, x_{ij}^{iter} is a value of decision variable j of current individual i at iteration $iter$, p_r is a parameter for controlling diversification and intensification, x_{aj}^{iter} and x_{bj}^{iter} are values of decision variable j of selected current individuals a and b at iteration $iter$ ($1 \leq a \neq b \leq \mathrm{N}$), N_I is the number of individuals, N_{DV} is the number of decision variables x_{ij}^{iter}, $rand(0,1)$ is an uniform random number in the range (0,1).

When two clusters are utilized for update, a new individual is updated using the following equation:

$$x_{ij}^{iter} = rand(0,1) \times x_{ij1}^{iter} + (1 - rand(0,1)) \times x_{ij2}^{iter}$$
$$(i = 1, \cdots, N_I, \quad j = 1, \cdots, N_{DV}) \quad (9)$$

where x_{ij1}^{iter}, and x_{ij2}^{iter} are selected current decision variable j of individual i at iteration $iter$ at each cluster.

- Step 4 The current individual is replaced with the newly generated individual when the newly generated individual is superior to the current one.

2) *Overview of PMP-MBSO*

Using PMP, several sub-populations are created from one whole population. A process is allocated to each sub-population and optimization procedures in each sub-population are calculated in parallel. Figure 8 shows an example of PMP-MBSO with four sub-populations. Each sub-population independently performs MBSO procedures using individuals in each sub-population [22]. When iteration number reaches a certain number, individuals are replaced among connected sub-populations. This action is called migration. Hyper parameters of PMP are migration policy, sub-population network topology,

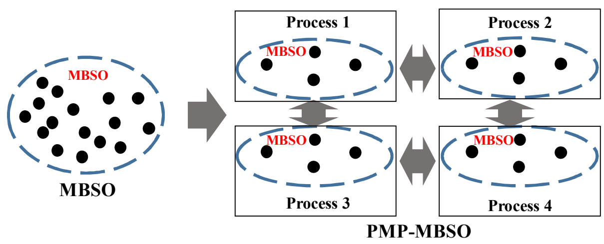


Fig. 8. An example of PMP-MBSO with four sub-populations.

Algorithm PMP-MBSO

Initializing: Individuals are initially generated at random within upper and lower limits.
for $s = 1\ to\ S$ ($the\ number\ of\ sub-population$), do:
  S Sub-populations are allocated to S processes for speed-up of computation time.
  for $iter = 1\ to\ itermax$, do:
    The following procedures are performed in parallel in each sub-population.
    Clustering: SGM is utilized for grouping all individuals into several clusters. At each cluster, a cluster center is changed from center of distance to the best evaluated individual.
    if ($rand(0,1) < p\_replace$), do
      One cluster center is replaced with a random individual.
    end of if
    for $i = 1\ to\ I$ (the number of individuals), do:
      if ($rand(0,1) < p\_one$), do:
        One cluster is randomly selected.
        if ($rand(0,1) < p\_r$), do:
          New individual is generated using (7).
        else if, do:
          if ($rand(0,1) < p\_one\_cluster$), do:
            An individual is randomly selected which is not the cluster center, and new individual is generated using (8).
          else if, do:
            An individual is randomly selected which is the cluster center, and new individual is generated using (8).
          end of if
        end of if
      else if, do:
        Two clusters are randomly selected.
        if ($rand(0,1) < p\_r$), do:
          New individual is generated using (7).
        else if, do:
          if ($rand(0,1) < p\_two\_cluster$), do:
            Two individuals are randomly selected which are not the cluster center, and new individual is generated using (8) and (9).
          else if, do:
            Two individuals are randomly selected which are the cluster center, and new individual is generated using (8) and (9).
          end of if
        end of if
      end of if
      If the newly generated individual is superior to the current one in each individual, the newly generated individual is utilized as the current one.
    end of for
    When a certain number of iterations is reached, individuals are replaced among connected sub-populations.
  end of for
end of for

Fig.9. An algorithm of PMP-MBSO.

migration interval, and the number of sub-populations. In the simulation, migration policy and migration interval are determined by pre-simulation. Various numbers of sub-populations and sub-population topologies are investigated in simulation. An algorithm of PMP-MBSO can be summarized in Fig. 9.

D. The proposed refrigerated showcase fault detection method by PANNs using PMP-MBSO and Correntropy

A conception of parameter training of PANNs using PMP-MBSO and Correntropy for the proposed fault detection method for refrigerated showcases is shown in Fig. 10. The algorithm of proposed fault detection is explained below.

Step 1 M independent training subsets are created from sensor data of refrigerated showcases. M Classifiers are created for M training subset.

Step 2 For all training subsets, ANN parameters of all individuals are initially generated within constraints.

Step 3 Using forward propagation, for all individuals, outputs of the ANN are calculated using each training subsets. Using Correntropy, a summation of LFVs is calculated using errors between target values and outputs of ANN. A summation of LFVs using Correntropy is applied to evaluate the ANN parameters of all individuals for all training subsets.

Step 4 For mth classifier, S sub-populations are created from one whole population of initial individuals. $iter$ is set to 1.

Step 5 S sub-populations are calculated using S processes for speed-up of training time. MBSO procedures including clustering and new individual generation are performed in each process for the mth training subset in parallel.

Step 6 Using newly generated ANN parameters, LFVs of individuals using Correntropy are calculated. The current individual is replaced with the newly generated individual in each sub-population for the mth training subset when the newly generated individual's LFV is superior to the current individual's LFV in each individual.

Step 7 Migration is performed at each sub-population by transferring individuals among sub-populations when iteration number reaches a certain number.

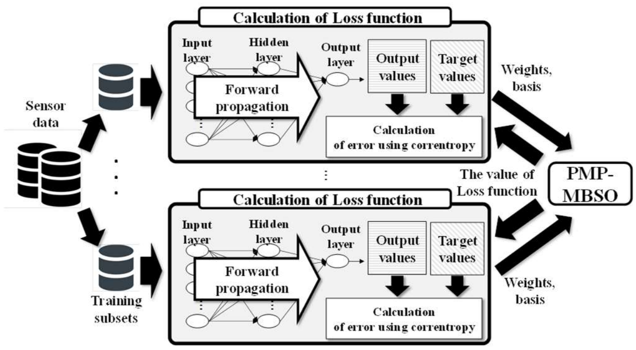


Fig. 10. A conception of parameter training of PANNs using PMP-MBSO and Correntropy for the proposed fault detection method for refrigerated showcases.

Step 8 If $iter$ reaches preset $itermax$, go to Step 9. Otherwise, $iter = iter + 1$ and go back to Step 5.

Step 9 Training procedures can be stopped and go to Step10 if all classifiers have been trained. Otherwise, $m = m + 1$ and move back to Step 4.

Step10 For all classifiers, using test data, output values are calculated using forward propagation. Final results are calculated with output values of each classifier using decision by majority.

V. SIMULATION

A. Simulation conditions

Using actual operation data of refrigerated showcase, the proposed PANNs with PMP-MBSO and Correntropy, conventional ANNs with SGD and LSE [7], and ANNs with MBSO and Correntropy [10] for fault detection of refrigerated showcase are compared. Table 1 shows general simulation conditions.

Table 2 shows common parameters of ANN using evolutionary computation techniques and Correntropy, and various parameters of each method. The best individual among all the best individuals sent by the sending sub-populations is utilized for replacement of the worst individual in the receiving sub-population using the W-B policy.

B. Simulation results

Table 3 shows average detection accuracy rates using two abnormal rates for test data by the ANN with SGD and LSE [7], the ANN with MBSO and Correntropy [10], and the proposed PANNs with PMP-MBSO and Correntropy. Using the proposed methods, especially using the trigonal pyramid topology, "Fault" rates are higher than those of the conventional methods in both cases of 0% and 10% of abnormal rates (bold

TABLE I. GENERAL SIMULATION CONDITIONS

General simulation conditions
- All operation data of refrigerated showcases are divided into 70% training and 30% test data.
- Since the number of normal data is exceedingly larger than the number of fault data, 99% of normal data are reduced using random under sampling.
- In the training data, 10% normal data are replaced with incorrectly measured data at random (abnormal rate).
- 10 units of hidden layer are utilized.
- As activation functions of ANN, Sigmoid functions are utilized.
- Since training of ANN depends on initial ANN parameters, the number of simulation trials with different random seeds is set to 30.
- Ratios to detect normal condition as normal correctly (Normal), ratios to detect abnormal condition as faulty correctly (Fault), and ratios of correct fault detection for whole test data (Total) are utilized for evaluating three methods.
- Since reduction of negative impact on customers by abnormal conditions is the most important for fault detection of refrigerated showcases, "Fault" evaluations are regarded as most important.

TABLE II. COMMON PARAMETERS OF ANN USING CORRENTROPY AND EVOLUTIONARY COMPUTATION TECHNIQUES, AND PARAMETERS OF EACH METHOD

Common parameters of ANN using Correntropy and evolutionary computation techniques	
Kernel size: 0.01, The number of individuals: 384, $itermax$: 500, $P\_replace$: 0.1, P_one: 0.9, P_one_center: 0.7, P_two_center: 0.9, the number of cluster: 4	
Parameters for each method	
Method	**Parameter Settings**
MBSO and PMP-MBSO	P_r: 0.001
PMP-MBSO	The number of sub-populations (sub-pops.): 1, 2, 4, and 8 Migration policy: Worst-Best policy Migration interval: 10 Sub-population network topology: Ring (2, 4, and 8 sub-pops.), Trigonal pyramid (4 sub-pops.), Cube (8 sub-pops.)
Pasting	The number of classifiers: 7
SGD	Learning rate: 0.01, T (Maximum iteration): 1000000

numbers). Consequently, even if incorrectly measured data exist in training data, abnormal status can be detected more precisely by the proposed method than the ANN with SGD and LSE, and the ANN with MBSO and Correntropy. Namely, deterioration of perishable products' quality can be prevented using the proposed methods the most among all methods.

Mean ranks and results of Friedman test using the ANN with SGD and LSE, the ANN with MBSO, and the proposed PANNs with PMP-MBSO and Correntropy through 30 trials are shown in Table 4. Using D'Agostino-Pearson and Anderson-Darling tests, since the results are verified to lack for regularity, Friedman test is applied. As observed in table 4, average ranks of the proposed PANNs with PMP-MBSO and Correntropy are superior to those of the conventional ANN with SGD and LSE, and ANN with MBSO and Correntropy. Therefore, superiority of the proposed method is verified with 0.05 significant level.

Figure 11 shows average calculation time by the proposed PANNs using PMP-MBSO and Correntropy using 1, 2, 4, and 8 processes though 30 trials. As observed in fig. 11, the proposed PANNs using PMP-MBSO and Correntropy using 8 processes is about 2.5 times faster than those using one process.

TABLE III. AVERAGE DETECTION ACCURACY RATES USING VARIOUS ABNORMAL RATES FOR TEST DATA BY THE CONVENTIONAL ANN WITH SGD AND LSE, THE CONVENTIONAL ANN WITH MBSO AND CORRENTROPY, AND THE PROPOSED PANNS WITH PMP-MBSO AND CORRENTROPY

Method		**Evaluation item**	**Average detection accuracy rates** [%]	
			0 [%] Abnormal rates	**10 [%] Abnormal rates**
ANN using SGD and LSE [7]		Total	96.361	94.117
		Normal	97.752	94.675
		Fault	95.531	92.651
ANN using MBSO and Correntropy [10]		Total	96.518	95.787
		Normal	96.198	96.130
		Fault	97.358	94.888
The proposed PANNs using PMP-MBSO and Correntropy	Num. of sub-population: 1	Total	96.198	95.858
		Normal	95.048	94.544
		Fault	99.215	99.307
	Num. of sub-population: 2	Total	96.401	95.975
		Normal	95.286	94.706
		Fault	99.329	99.307
	Num. of sub-population: 4 Topology: Ring	Total	96.661	96.296
		Normal	95.661	95.141
		Fault	99.284	99.329
	Num. of sub-population: 4 Topology: Trigonal pyramid	Total	96.210	95.784
		Normal	95.021	94.433
		Fault	**99.329**	**99.329**
	Num. of sub-population: 8 Topology: Ring	Total	96.556	96.444
		Normal	95.533	95.354
		Fault	99.239	99.307
	Num. of sub-population: 8 Topology: Cube	Total	96.463	95.994
		Normal	95.379	94.723
		Fault	99.306	99.329

TABLE IV. MEAN RANKS AND RESULTS OF FRIEDMAN TEST THROUGH 30 TRIALS USING THE ANN WITH SGD AND LSE, THE ANN WITH MBSO AND CORRENTROPY, AND THE PROPOSED PANNS WITH PMP-MBSO AND CORRENTROPY

Abnormal rates	**ANN using SGD and LSE**	**ANN using MBSO and Correntropy**	**The proposed PANNs using PMP-MBSO**	***P*_value**
0%	2.67	2.33	1	7.35E-11
10%	2.97	2.03	1	2.46E-13

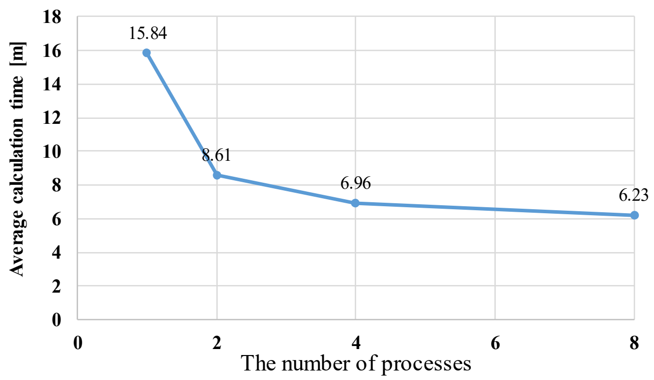


Fig. 11 Average calculation time by the proposed PANNs using PMP-MBSO and Correntropy using 1, 2, 4, and 8 processes though 30 trials.

Consequently, the proposed method can be verified to reduce training time.

VI. Conclusions

This paper proposes a refrigerated showcases fault detection method by Pasting based Artificial Neural Networks (PANNs) using Parallel Multi-population Modified Brain Storm Optimization (PMP-MBSO) and Correntropy. Verification of effectiveness of the proposed method is performed by comparison with the conventional ANN with SGD and LSE, and the conventional ANN with MBSO and Correntropy using actual operation data of refrigerated showcases. Even if incorrectly measured data exist in training data, abnormal status can be detected more precisely by the proposed method comparing with the ANN with SGD and LSE, and the ANN with MBSO and Correntropy. Furthermore, the proposed method can be verified to reduce training time. The characteristics of the proposed method are suitable for practical application.

To improve accuracy of fault detection and reduce training time, applications of novel training techniques for ANN parameters and other ensemble learning techniques will be investigated as future works.

References

[1] H. Wang, et al, "A study on fault detection and diagnosis for VAV air handling units of real buildings," Proc. of International Conference on Electric Technology and Civil Engineering (ICETCE), April 2011.

[2] K. Bruton, et al, "Results from testing of a “cloud based” automated fault detection and diagnosis tool for AHU's," Proc. of IEEE 18th Conference on Emerging Technologies & Factory Automation (ETFA), Sept. 2013.

[3] L. Chang, et al, "Cloud-Based parallel implementation of an intelligent classification algorithm for fault detection and diagnosis of HVAC systems," Proc. of Smart Cities Conference, Sept. 2017.

[4] H. Han, et al, "PCA-SVM-Based Automated Fault Detection and Diagnosis (AFDD) for Vapor-Compression Refrigeration Systems," HVAC&R Research, Vol. 16, pp.295~313, May 2010.

[5] Ministry of Economy, Trade and Industry, Preliminary Report on the Current Survey of Commerce, "http://www.meti.go.jp/statistics/tyo/syoudou/result/pdf/201908S.pdf (in Japanese).

[6] A. Santana, et al., “Machine learning application for refrigeration showcase fault discrimination, “ Proc. of IEEE Region Ten Conference (TENCON) 2016, Nov. 2016.

[7] A. Santana, et al., “One Class Data mining Approaches for Leakage Fault Detection in Refrigeration Showcases, “ Annual conference of Society of Instrument and Control Engineers (SICE), Sept. 2018.

[8] N. Otaka, et al., “Refrigerated Showcase Fault analysis by a Correntropy based Artificial Neural Network Using Fast Brain Swarm Optimization “, Proc. of ICSI, July 2019.

[9] N. Otaka, et al., “Refrigerated Showcase Fault analysis by a Correntropy based Artificial Neural Network Using Particle Swarm Optimization “, Annual conference of SICE, Sept. 2018.

[10] N. Otaka, et al., “Refrigerated Showcase Fault analysis by a Correntropy based Artificial Neural Network Using Modified Brain Swarm Optimization “, Proc. of SSCI, Dec. 2019.

[11] R. J. Bessa, V. Miranda, and J. Gama: "Entropy and Correntropy Against Minimum Square Error in Offline and Online Three-Day Ahead Wind Power Forecasting," IEEE Transactions on Power Systems, Vol. 24, No. 4, pp.1657-1666, Nov. 2009.

[12] Z. Wang, B. Wang, C. Liu, and W. Wang: "Improved BP neural network algorithm to wind power forecast," The Journal of Engineering, Vol. 2017, Issue 13, pp. 940-943, 2017.

[13] D. Sakurai, Y. Fukuyama, et al., “Daily Peak Load Forecasting by Artificial Neural Network using Differential Evolutionary Particle Swarm Optimization Considering Outliers, “ Proc. of Information Federation of Automatic Control (IFAC) Control of Smart Grid and Renewable Energy Systems (CSGRES) 2019, June 2019.

[14] C. Zhang, et al., “Particle swarm optimization for evolving artificial neural network, “ Proc. of IEEE International Conference on Systems, Man, and Cybernetics (SMC2000) , Oct. 2000.

[15] Z. Zhan, Y. Shi, et al., “A Modified Brain Storm Optimization,” Proc. of the IEEE World Congress on Computational Intelligence , June 2012.

[16] M. Sato, Y. Fukuyama, et al., “Total Optimization of Smart City by Modified Brain Storm Optimization, “ Proc. of IFAC Control of Power and Energy System (CPES) 2018, Sept. 2018.

[17] K. Arai, Y. Fukuyama, et al., “Optimal Operational Planning of Energy Plants by Modified Brain Storm Optimization, “ Proc. of IFAC CPES 2018, Oct. 2018.

[18] D. Azuma, Y. Fukuyama, et al., “Modified Brain Storm Optimization for Load Adjustment Distribution State Estimation Using Correntropy, “ Proc. of IEEE TENCON 2018, Oct. 2018.

[19] L. Breiman, “Pasting Small Votes for Classification in Large Databases and On-Line,” Machine Learning, vol. 36, pp. 85-103, July 1999.

[20] N. V.Chawla, et al., “Learning Ensembles from Bites: A Scalable and Accurate Approach,” Journal of Machine Learning Research, Vol. 5, pp.421~451, 2004.

[21] G. Valentini, “An experimental bias-variance analysis of SVM ensembles based on resampling techniques,” IEEE Transaction on Systems, Man, and Cybernetics, Vol. 35, No. 6, Dec, 2005.

[22] D. Azuma, Y. Fukuyama, et al., “Dependable Parallel Multi-population Modified Brain Storm Optimization for Distribution State Estimation considering Outliers using Just-in-time Modeling and Correntropy,” IEEJ Transaction on Power and Energy, Vol.140, No.2, Feb. 2020 (to appear in Japanese).

[23] D. Rumelhart, et al., Parallel Distriduted Processing, MIT press, 1986.

[24] W. Liu, P. P. Pokharel and J. C. Principe, “Correntropy: A Localized Similarity Measure,” Proc. of International Joint Conference on Neural networks (IJCNN) , July 2006.

[25] Y. Shi, et al., “Brain Storm Optimization Algorithm: a review,” Artificial Intelligence Review, Vol.46, Issue 4, pp.445-458, Feb. 2016.