energies

MDPI

Article

Identification of Inrush Current Using a GSA-BP Network

Zhou Ruhan, Nurulafiqah Nadzirah Binti Mansor * and Hazlee Azil Illias

Electrical Engineering, University Malaya, Kuala Lumpur 50603, Malaysia
* Correspondence: afiqah.mansor@um.edu.my; Tel.: +60-1-9427-6319

Abstract: Ensuring a stable and efficient transformer operation is a very crucial task nowadays, especially with the integration of modern and sensitive electrical equipment and appliances down the line. However, transformer maloperation still cannot be completely avoided, particularly with the existence of inrush current that possess similar characteristics as the fault currents when a fault occurred. Thus, this paper proposes an enhanced method for inrush current identification based on a backpropagation (BP) network, optimized using genetic and simulated annealing algorithms. The proposed method has the ability to find the global optimal solution while avoiding local optima, with increased solution accuracy and low calculation complexity. Through extensive simulations, it was found that the inrush and fault currents have differences in their harmonic contents, which can be exploited for the identification of those currents using the proposed identification method. The proposed genetic simulated annealing–BP (GSA-BP) algorithm make use of 200 current samples to improve the detection accuracy of the inrush current from 80% to 97.5%. Comparative studies performed against the existing identification methods show that the GSA-BP network has superior efficiency and accuracy while being practical for real-life application to improve the transformer protection system.

Keywords: inrush current; harmonic components; BP network; genetic algorithm; simulated annealing algorithm

Citation: Ruhan, Z.; Mansor, N.N.B.; Illias, H.A. Identification of Inrush Current Using a GSA-BP Network. *Energies* **2023**, *16*, 2340. https://doi.org/10.3390/en16052340

Academic Editor: Alberto Geri

Received: 26 November 2022
Revised: 3 February 2023
Accepted: 22 February 2023
Published: 28 February 2023

1. Introduction

Differential protection is a widely used approach for power transformer protection, the performance of which is constantly challenged by the inrush current [1]. Therefore, precise differentiation between the inrush and fault currents for a split second is crucial for the operation of a transformer [2,3].

Although conventional inrush current detectors, such as second harmonic criteria and gap detectors, can block the influence of inrush current in most scenarios, a lack of precision in common detection is inevitable [4–7]. These classifications have been integrated into new methods to improve reliability. However, maloperation still occurs in certain scenarios, such as energizing with a high initial flux [8–10]. The inrush current is twice or ten times the normal rated current. Such unnecessary current spikes can damage equipment like transformers. The inrush current can cause the circuit breaker to trip because it has a similar amount of current as the fault current, causing the transformer protection device to mistake the inrush current for a fault and cut the circuit. Therefore, reliable identification of the inrush current is highly desired. The method proposed in this paper is to reduce or even avoid these situations as much as possible. To enhance the reliability of differential protection [11–13], researchers have developed several techniques over the past decades to address the inrush problem. The well-known methods are as follows:

- The discontinuous angle principle [9]: discontinuous angle principle applied to quickly recognize the inrush current and fault current. When the discontinuous angle of current is greater than 65°, it is determined that the transformer inrush current occurs. The transformer protection may wrongly determine that short-circuit current is caused

by a power component fault. Discontinuous angle principle aims to avoid tripping when the inrush current occurs. If the wave width of the current is larger than 140° and the discontinuity angle is less than 65°, this current will be identified as a fault current and the relay should quickly switch off. Once the discontinuity angle is greater than 65°, the relay should be keep locked. However, reverse current will occur when the current transformer saturates. It will make the discontinuous angle of the inrush current disappear, which means the protection will wrongly operate. When the internal fault occurs, saturated current transformer causes the discontinuous angle of differential current to increase. The protection will lose efficacy. Moreover, accurately measuring the discontinuity angle remains a difficult problem to overcome. When an internal fault occurs, transient high harmonics distort the current waveform. The waveform distortion does not cause discontinuous angle, but it affects the wave width of the current. If the waveform distortion is so severe that the wave width is less than 140°, it will cause a delay in protection or even a wrong judgment;

- Second harmonic braking [12,14]: inrush current contains much higher harmonic components, especially the second harmonic components. According to this characteristic, the second harmonic brake was proposed. Once the protective device detects the value of the second harmonic in a differential current being less than the setting value, the protective relay will be latched to prevent the maloperation.

The amplitude ratio of the second harmonic and fundamental harmonic is the second harmonic braking ratio. In a practical application, the inrush current that has been already ignored by differential protection in operating experience and no-load closing experiments is selected. The lowest harmonic content of this inrush current can be a setting value.

By detecting the magnitude of the second harmonic content in the three-phase current, it is judged whether it is an inrush current or not. The criterion of the inrush current is

$$I_{d2} > K_2 I_{d1} \tag{1}$$

This method is simple and practical, and there are many valuable experiences in the actual operation, most of the transformer microcomputer protection adopts this principle. However, in practical applications, the second harmonic discrimination principle still faces many problems. The selection of the second harmonic braking ratio K_2 is more difficult and there is no uniform standard.

In the power system installed with static reactive power compensation devices or other capacitive components, the fault current also has a large second harmonic content. Identification speed and even the accuracy of the second harmonic braking can be affected.

When fault happens in the transformer before no-load closing, the fault current is in the fault term after closing and the inrush current is in the non-fault term. Differential protection must be on work in this situation, but second harmonic braking may classify it as an inrush current and keep the protection locked, and the protection cannot operate normally. In addition, second harmonic decay is very slow and this may cause a long action time.

- Wavelet theory [15–17]: the original form of most signals is a time-domain signal, which means the signal is always a function about time. The most useful information is usually hidden in the frequency content of the signal. Waveform theory is used to search such information. The wavelet transform is performed on the time-domain signal, and the frequency spectrum of the signal will be obtained. It is applied in power systems for predicting the efficiency of solar power and analyzing [18] the weak signal under noise interference [19]. It can transform the waveforms of the current, and makes the difference between the inrush current and fault current be better reflected. The mutation part of the signal and the singular signal always contain important information. Wavelet analysis possesses a local refinement on the current signal, which can satisfy requirements both in the time and frequency domain. When a transformer fault occurs, the current signal in the operating power system will present

singularity. Waveform analysis is able to complete a detailed analysis of current signals at different scales. The components of each scale correspond to different frequency components in the current signal. The results of wavelet analysis reveal the distortion points of the current signal. The signal can receive a different local maximum by varying the waveform at different scales. By observing the trend of the maximum, the purpose of identifying the type of current can be achieved.

The waveform transform results of the inrush current show distortion in every cycle periodically. The result of the fault current shows a large distortion in a short time when fault occurs; then, the waveform is flat and almost with no distortion. These are the most obvious characteristics of inrush and fault currents in the waveform transform identification. However, it requires long data windows and is susceptible to noise and unanticipated disturbances.

- Multi-criteria decision-making: MCDM is a branch of operations research and it is used to balance and evaluate multiple conflicting goals for obtaining a satisfactory result. The most common application is the evaluation of costs and prices. This method has been applied in power systems such as finding suitable places for solar power plants [20]. It can also be used for transformer protection [12]. Compared with the independent decision-making of the original protection of the transformer, it can cover the blind spots of the protection, such as the restricted earth fault relay being sensitive in low-level internal earth faults (the faults occurring within 20–30% of the winding near the neutral point) causing the transformer protection relay to not work and making the protection more reliable. However, part of the indicators that are needed to be balanced are vague and cannot be quantified, such as the effects of weather and the aging of materials. Additionally, there is no common measure or unit of measure for goals, making them difficult to be compared.

A reliable classifier should have accuracy and efficiency. It is essential to maintain a high computational efficiency in fresh scenarios, as it is continuously restrained by various factors [21,22]. Feasibility is also one of the considerations. The genetic simulated annealing–backpropagation (GSA-BP) network satisfies all these requirements.

In this paper, an enhanced method based on a GSA-BP model and waveform is proposed to discriminate the inrush current from the fault current in transformers. First, the waveforms of the inrush and fault currents were simulated in MATLAB/Simulink. The classification model consists of a BP network and genetic algorithm (GA). Finally, the fundamental wave and second to fifth harmonics of the inrush and fault currents collected using fast Fourier transform (FFT) analysis were used as the input of the GSA-BP network. Numerous training and verification simulations were performed on the GSA-BP network. Furthermore, the performance of the GSA-BP network was compared with that of the BP network. GSA-BP can solve the problem of low accuracy of existing transformer protection methods, and also improves the recognition speed. The results from the simple three-phase power circuit and IEEE 14 bus test system show that the recognition rate of GSA-BP for current samples reached 97.5%. In addition, it only takes 0.186 s to recognize a sample on average. Compared with the original BP network accuracy, the accuracy rate has increased by 17.5%. Compared with the proposed optimized BP network, such as the GA-BP recognition method, the accuracy and operating efficiency have increased by 9.5% and 99.3%, respectively. The proposed GSA-BP model can successfully classify the inrush and fault currents more quickly with promising accuracy.

The remainder of this paper is organized as follows:

Section 2: The common inrush current identification and existing problems.

Section 3: Simulation and application of the proposed GSA-BP network.

Section 4: Discussion and summary of performance of the GSA-BP method.

2. Methodology

2.1. Introduction to BP Neural Network

A BP network is a multilayer neural network that can train the weights of nonlinear differentiable functions.

The basic model of the neurons in the BP neural network is shown in Figure 1. The inputs are connected to the next layer of neurons through the corresponding weights w, and the final outputs can be expressed by the following formula:

$$a = f(wp + b) \tag{2}$$

f is a function of inputs and outputs, w is the connection weight between neurons, and b is the bias.

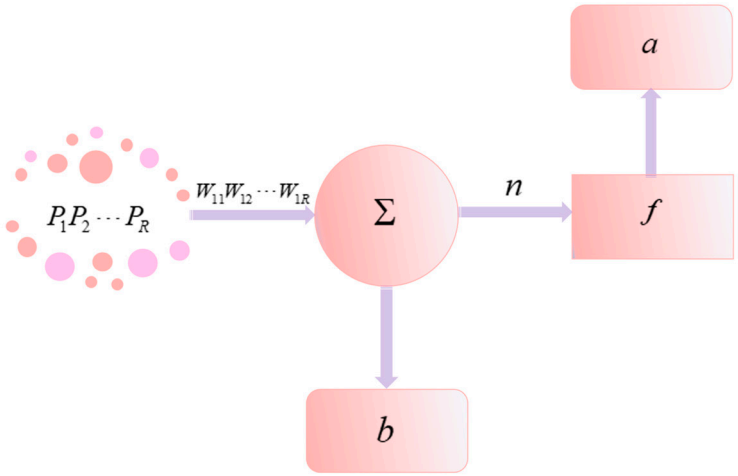


Figure 1. Single neuron of a BP network.

The structure of the BP network is shown in Figure 2. The number of inputs is R, and the hidden layer has only one layer with S neurons. Each neuron is a type of sigmoid.

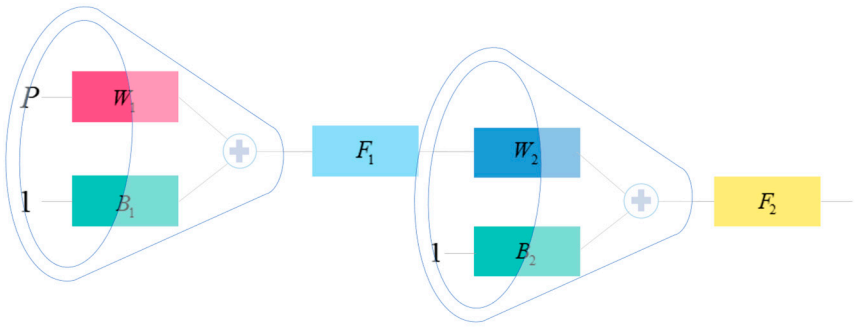


Figure 2. BP network structure.

2.2. Learning Rules

The BP network algorithm is a supervised δ algorithm. Its principle is to set the learning samples as inputs and corresponding outputs. The learning process compares the actual outputs of the BP network with the expected outputs. The quadratic sum of their errors are used to modify the weight of the neuron in reverse. Finally, the quadratic sum of the error between the actual and desired outputs reaches a certain range. When the error is large, its influence on the weight also increases.

The algorithm derivation process of the BP neural network is as follows: If P is the input of the network, there are R input neurons, s1 is the number of hidden layer neurons, and s2 is the number of output neurons. F1 is the hidden layer activation function, f_2 is the output layer activation function, A is the actual output, and T is the expected output.

The forward propagation process of information is represented by Equations (3)–(5). The output of the i-th neuron in the hidden layer is

$$a1_i = f1(\sum_{j=1}^{r} w1_{ij}p_j + b1_i)(i = 1, 2 \ldots, s1) \tag{3}$$

The output value of the kth neuron in the output layer is

$$a2_k = f2(\sum_{i=1}^{s1} w2_{ki}a1_i + b2_k)(k = 1, 2, \ldots, s2) \tag{4}$$

The error of the network is expressed as follows:

$$E(W, B) = \frac{1}{2}\sum_{k=1}^{s2} (t_k - a2_k)^2 \tag{5}$$

The process of neuron weight modification and error reverse transmission is as follows: The weights from the i-th input to the k-th output are

$$\begin{aligned} \Delta\omega 2_{ki} &= -\eta\frac{\partial E}{\partial \omega 2_{ki}} = -\eta\frac{\partial E}{\partial a2_k}\cdot\frac{\partial a2_k}{\partial \omega 2_{ki}} \\ &= \eta(t_k - a2_k)\cdot f2'\cdot a1_i = \eta\cdot\delta_{ki}\cdot a1_i \end{aligned} \tag{6}$$

$$\delta_{ki} = (t_k - a2_k)\cdot f2' = e_k\cdot f2' \tag{7}$$

$$e_k = t_k - a2_k \tag{8}$$

From this, it can be derived

$$\begin{aligned} \Delta b2_{ki} &= -\eta\frac{\partial E}{\partial b2_{ki}} = -\eta\frac{\partial E}{\partial a2_k}\cdot\frac{\partial a2_k}{\partial b2_{ki}} \\ &= \eta(t_k - a2_k)\cdot f2' = \eta\cdot\delta_{ki} \end{aligned} \tag{9}$$

The process of using the gradient descent method to modify the weights of the hidden layer neurons, for the weights from the i-th input to the k-th output, is as follows:

$$\begin{aligned} \Delta\omega 1_{ij} &= -\eta\frac{\partial E}{\partial \omega 1_{ij}} = -\eta\frac{\partial E}{\partial a2_k}\cdot\frac{\partial a2_k}{\partial a1_i}\cdot\frac{\partial a1_i}{\partial \omega 1_{ij}} \\ &= \eta\sum_{k=1}^{s2}(t_k - a2_k)\cdot f2'\cdot\omega 2_{ki}\cdot f1'\cdot p_j = \eta\cdot\delta_{ij}\cdot p_j \end{aligned} \tag{10}$$

$$\delta_{ij} = e_i\cdot f1', e_i = \sum_{k=1}^{s1} w2_{ki}\delta_{ki}, \delta_{ki} = e_k\cdot f2', e_k = t_k - a2_k \tag{11}$$

Therefore,

$$\Delta b1_i = \eta\delta_{ij} \tag{12}$$

2.3. Introduction to Genetic and Simulated Annealing Algorithms

A GA is a global search heuristic algorithm used to solve optimisation problems in the field of artificial intelligence. It overcomes the limitations of conventional algorithms, which can easily fall into local minima. GA is originally developed based on evolutionary phenomena in biology, including heredity, mutation, natural selection, and hybridization [23–25].

The GA is a search algorithm with an iterative process of 'generate and test'. First, solutions to the problem are encoded to function as suitable expressions for genetic operations. Subsequently, the fitness function is established based on the objective function of the optimisation problem. When the fitness function is certain, natural selection determines the

chromosomes that are suitable for survival based on the fitness value. Finally, the surviving individuals form a population that can be reproduced (Figure 3).

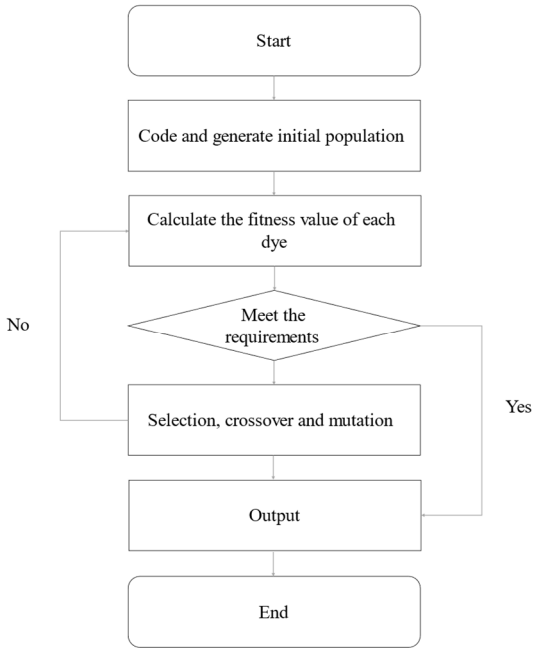


Figure 3. Genetic algorithm flow chart.

2.4. Design of the GSA-BP Algorithm

2.4.1. Coding

The BP network includes connection weights between input layer and implicit layer, implicit layer threshold, connection weights between implicit layer and output layer, and output layer threshold. When GSA-BP classifies currents, it is necessary to find out the appropriate weights and thresholds that can guarantee the accuracy of the current classification. GA has the most important role in figuring out this problem.

The first stop of the GA-optimizing neural network is to represent the weights and thresholds as strings of real numbers. Each string is a chromosome. [26–29]

The parameter gene adopts real number encoding. The weight and threshold of the network can be negative, and the learning rate must be positive. Therefore, parameter genes should be divided into two parts: weight and rate genes. The coding length of the weighted gene is equal to the total number of weights and thresholds in each layer of the neural network. The coding length of the rate gene is 1. The coding length of the parameter gene is the sum of the length of the weight and rate genes, and the total coding length of a chromosome is the sum of the coding length of the connecting and parameter genes.

2.4.2. Decoding

Assuming that 1101 . . . 1 is the code of the connecting gene in a chromosome and the length of the code is K, the structure of such a three-layer BP neural network is shown in Figure 4. In the figure, i_n is the input layer neuron, h_k is the hidden layer neuron, o_m is the output layer neuron, X_n is the input sample, and Y_m is the network output. During decoding, the weights $\omega_{n,k}$ and $\omega_{k,m}$ from the input layer to the hidden layer and from the urgent layer to the output layer, and the value θ_k of the hidden layer are shown in Equations (13)–(15).

$$\omega_{n,k} = [\omega_{1,1}\omega_{1,2}0\omega_{1,4} \ldots \omega_{1,K}\omega_{2,1}\omega_{2,2}0\omega_{2,4} \ldots \omega_{2,K} \ldots\ldots \omega_{N,1}\omega_{N,2}0\omega_{N,4} \ldots \omega_{N,K}] \quad (13)$$

$$\omega_{k,m} = [\omega_{1,1}\omega_{1,2}\ldots\omega_{1,M}\omega_{2,1}\omega_{2,2}\ldots\omega_{2,M}00\ldots0\omega_{4,1}\omega_{4,2}\ldots\omega_{4,M}\ldots\ldots\omega_{K,1}\omega_{K,2}\ldots\omega_{K,M}] \quad (14)$$

$$\theta_k = [\theta_1\theta_20\theta_4\ldots\theta_K] \quad (15)$$

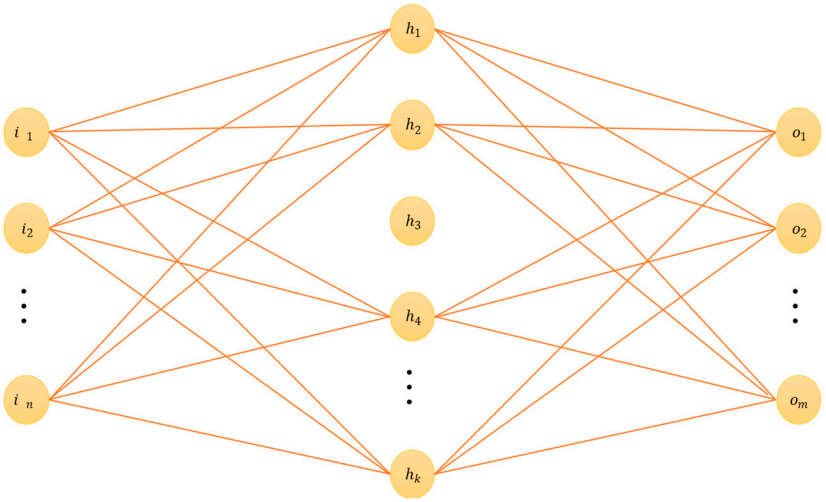


Figure 4. Neural network structure.

The learning rate has no effect on the network structure. It does not participate in the decoding.

The connection gene is binary-coded, and the number of nodes in the hidden layer is the same as that in the code string. The network structure is clear after decoding. The parameter genes are encoded in real numbers. The weights, thresholds, and learning rates of the network are obtained without complicated decoding.

2.4.3. Fitness Function

In the GA, the fitness function value is used to measure the closeness of each individual to the optimal solution in the optimisation calculation. The fitness function should generally satisfy two requirements: it should not be negative and it should be as simple as possible to reduce the complexity of the calculation [29–32].

The fitness function of the GA-BP algorithm is based on the total error of the neural network; that is, the fitness function of each chromosome is

$$f = 1/(1+E) \quad (16)$$

$$E = \frac{1}{2}\sum_{k=1}^{K}\sum_{j=1}^{M}\left(T_j^k - Y_j^k\right)^2 \quad (17)$$

where E is the total error in the neural network, T_j^k is the ideal output, Y_j^k is the real output, and K is the number of samples.

Higher fitness means that the individual is more suitable for the current environment. From the perspective of current identification, the current samples left with high fitness will be very close to the conditions that have been set, making it more accurate to distinguish the current type.

2.4.4. Genetic Operator

Selection: The selection operation of the GSA-BP algorithm combines the optimal individual preservation and fitness ratio. Optimal individual preservation is implemented on the chromosomes in the current population. In this method, individuals with the highest

adaptability do not participate in crossover and mutation and are directly replicated to the next generation [33]. After optimal individual preservation, a fitness ratio is adopted for the contemporary group. Assuming that the population size is n, the probability of individual i being selected is

$$P_{s_i} = f_i \Big/ \sum_{j=1}^{n} f_J \tag{18}$$

In this formula, f_i represents the fitness function value of the first chromosome. In this method, the probability of each chromosome being selected is P_{s_i}, which reflects the proportion of an individual's fitness to the total individual fitness.

The combination of optimal individual preservation and fitness ratio enables individuals with high fitness to be selected directly for the next generation. Individuals with low fitness also have chances. The diversity of individuals can ensure that their fitness is close to the optimal solution and prevents the algorithm from falling into the local optimum.

Crossover: The one-point crossover is also called the simple crossover [34]. A crossover point is randomly set among individuals. When crossover is performed, the partial structures of the two individuals before or after the point are exchanged to generate two new individuals.

Arithmetic crossover generates a new individual from a linear combination of two individuals. If an arithmetic crossover is performed between two individuals, x_1 and x_2, then the two new individuals x' and x'_2 after the crossover are

$$\begin{cases} x'_1 = \alpha x_1 + (1-\alpha)x_2 \\ x'_2 = \alpha x_2 + (1-\alpha)x_1 \end{cases} \tag{19}$$

Among them, α is a real number of 0–1.

Mutation: The basic mutation randomly selects one or more loci in the individual coding sequence and changes the gene in these loci based on a probability value.

Uneven mutation is a random perturbation of the original gene, and the result after the perturbation becomes the new gene after mutation. The mutation operation is performed on each locus with the same probability.

Crossover and mutation work simultaneously to ensure the diversity of individuals. Assuming that there is only crossover, the new solution generated in the iterative process can only come from the existing current samples in the initial generation. If the key feature for constructing the optimal solution is missing in the initial population, the optimal solution cannot be obtained only by crossover.

2.4.5. Annealing

The annealing algorithm is used to optimise and adjust the new individuals generated by the GA, and the results are considered the individuals of the next generation group. The annealing algorithm adds the optimal local solutions as new individuals to the GA and uses them to adjust the search range. The cooling method adopted in this paper is $T = T_0 \cdot 0.98^{\mathrm{i}}$. T_0 is the initial temperature and i is the times of evolutions.

2.5. Operating Principles of GSA-BP Network

Currently, to select the appropriate population size N, crossover probability P_c, and mutation probability P_m, many scholars have conducted research, but the settings of these parameters have not yet been unified. The optimal parameter ranges recommended by Schaffer are $N = 20 \sim 30$, $P_c = 0.75 \sim 0.95$, and $P_m = 0.005 \sim 0.01$. The optimal parameter ranges suggested by Xi Yugeng are $N = 20 \sim 200$, $P_c = 0.5 \sim 1.0$, and $P_m = 0.0 \sim 0.05$.

$$P_c = \begin{cases} (f_{\max} - f')/(f_{\max} - f_v), & (f' \geq f_v) \\ 1.0, & (f' < f_v) \end{cases} \tag{20}$$

$$P_m = \begin{cases} 0.5 * (f_{\max} - f)/(f_{\max} - f_v), & (f \geq f_v) \\ (f_v - f)/(f_v - f_{\min}), & (f < f_v) \end{cases} \quad (21)$$

When the fitness values of individuals in the population tend to be optimal, P_c and P_m increase, and when the population fitness values are relatively scattered, P_c and P_m decrease. When P_c and P_m are low, individuals with a fitness value higher than the average fitness value of the group can be protected into the next generation. When P_c and P_m are high, individuals with a fitness value lower than the average fitness value of the group will be eliminated. Therefore, the GA ensures convergence while maintaining population diversity (notation table is in Appendix A).

The formula to determine the acceptance probability of changing from the current solution X_1 to the new solution X_2 using the Metropolis criterion is [28]

$$P = \begin{cases} 1, f(X_1) \geq f(X_2) \\ \exp[f(X_1) - f(X_2)/T_i], f(X_1) < f(X_2) \end{cases} \quad (22)$$

The GSA-BP algorithm steps are shown in Figure 5:

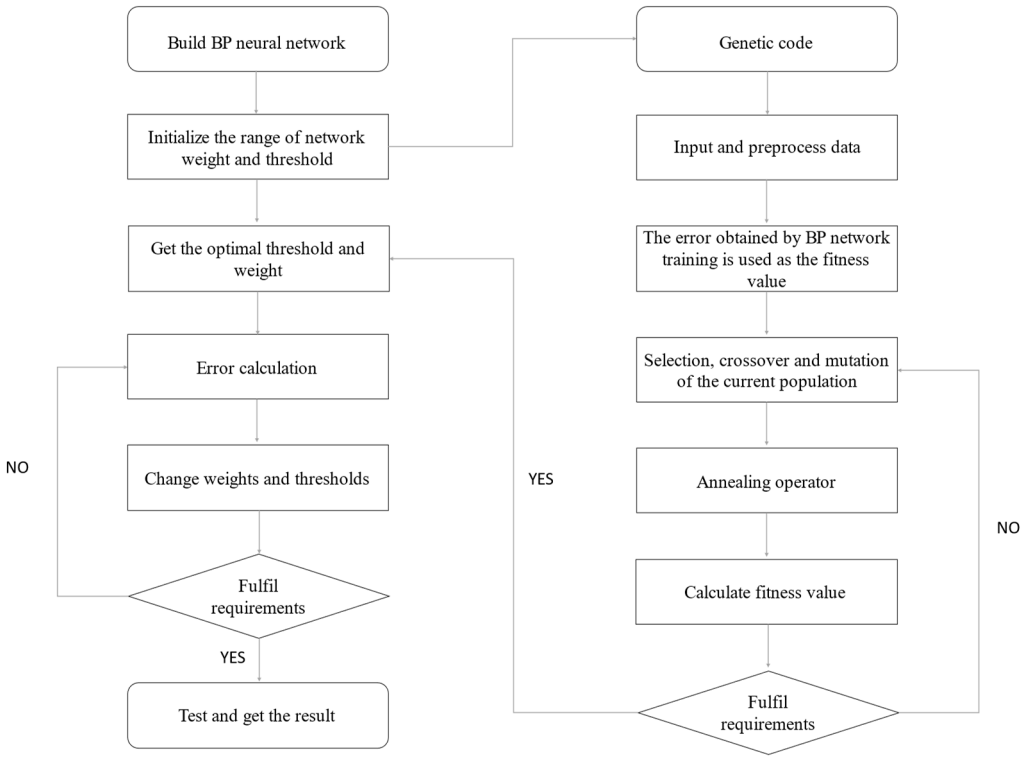


Figure 5. Training process of the GSA-BP network.

3. Results and Discussion

3.1. Simulation Model of Magnetizing Inrush Current

This section describes a simple power system with a dual-sided power supply and dual-winding transformer. The model parameter settings and simulation circuit are presented in Figure 6.

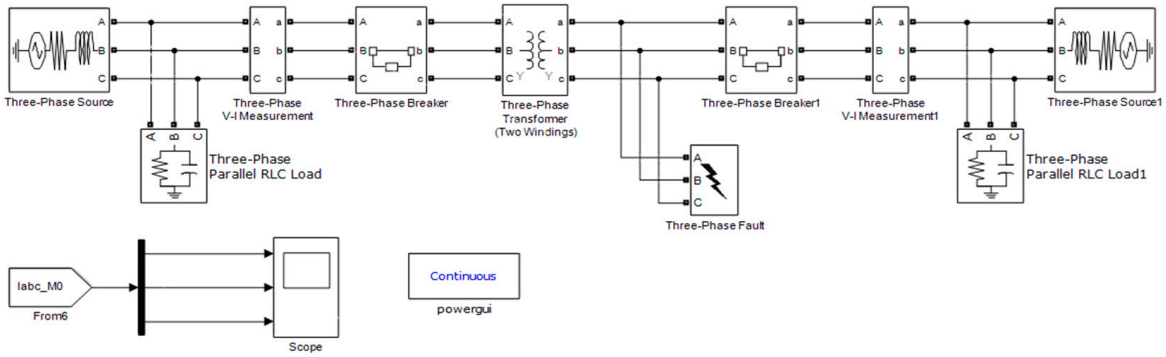


Figure 6. Transformer simulation.

The power component models used to build the simulation circuit in this study were selected from a power system module library. Table 1 provides a brief introduction of the components used.

Table 1. Waveform information of inrush current.

Initial Phase Angle/°	Residual Flux/pu	Peak Value/mA
0	0.5	1347
0	0.64	1471
0	0.8	1629
45	0.5	1274
45	0.64	1443
45	0.8	1609
90	0.5	651
90	0.64	778
90	0.8	924

3.2. Simulation of No-Load Energized

Inrush current is the power surge at the input of a power supply. It occurs during a short time period, when the input capacitors charge up after the input voltage is applied. In order to obtain the inrush current sample, the simulated circuit is required to be built in a no-load energized state. The frequency of the power source was 50 Hz and its voltage was 35 kV. The wiring form of the transformer winding was $Yd11$. In order to analyse the effect of the switch angle on the inrush current, the voltage switched at 0°, 45°, and 90° [4]. Meanwhile, three residual fluxes, 0.5, 0.64, and 0.8, were set in every switching under different angles [35].

In Figure 7, the variety of the initial phase angle not only affects the peak value of the current waveform, but also effects the form of the current waveform. Some information can be obtained from the graph. It presents that residual flux change will cause the difference of the peak value of the current. A small switching angle will lead to an intense inrush current. Additionally, the large residual flux has the same effect on the inrush current.

3.3. Simulation of the Fault Current

In order to find the common points and differences between inrush and fault currents, fault currents were simulated. This study on the fault current is based on three-phase transformers and simulates three common faults that may occur on them [36].

Figures 7 and 8 present both fault and inrush currents being much larger than normal. Table 2 details the data in Figure 8. The large magnitude of the current will cause a differential protective system, which should operate rapidly when faults occur and should not be interrupted by the inrush current. This is what they have in common and why a reliable method is required to differentiate them. Figure 9 shows that the waveform of the

inrush current is not a continuous waveform, and there are dead angles in every circle, while the fault current is continuous and close to a sine wave.

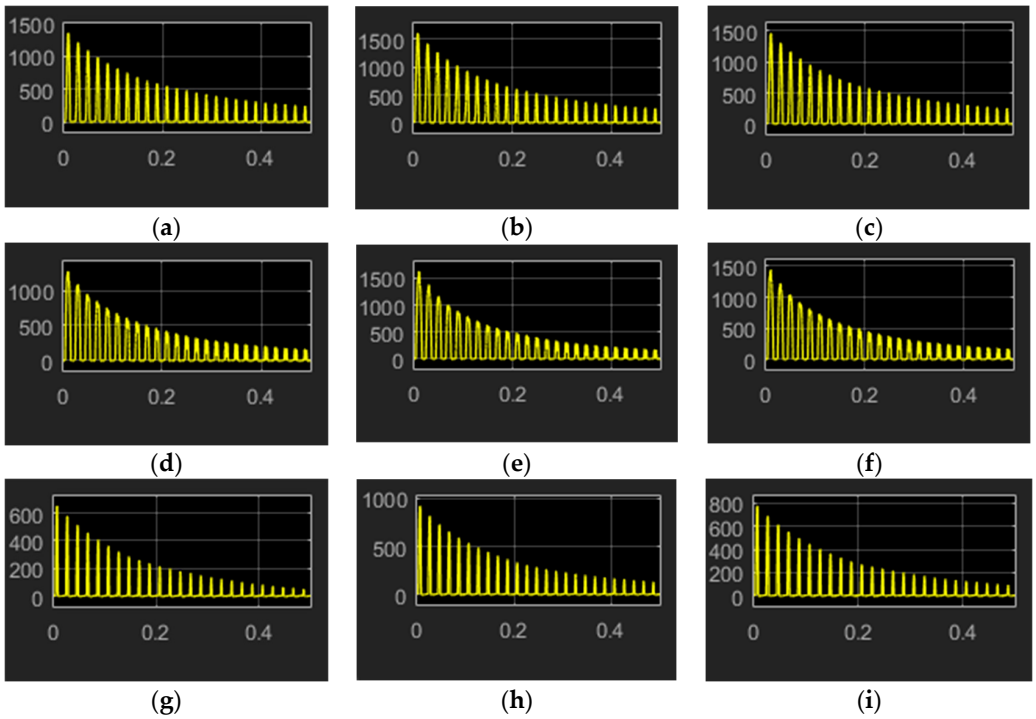


Figure 7. Inrush current under no-load energized with different initial phase angle and residual flux: (**a**) 0°, 0.5 pu; (**b**) 0°, 0.8 pu; (**c**) 0°, 0.64 pu; (**d**) 45°, 0.5 pu; (**e**) 45°, 0.8 pu; (**f**) 45°, 0.64 pu; (**g**) 90°, 0.5 pu; (**h**) 0°, 0.8 pu; (**i**) 90°, 0.64 pu.

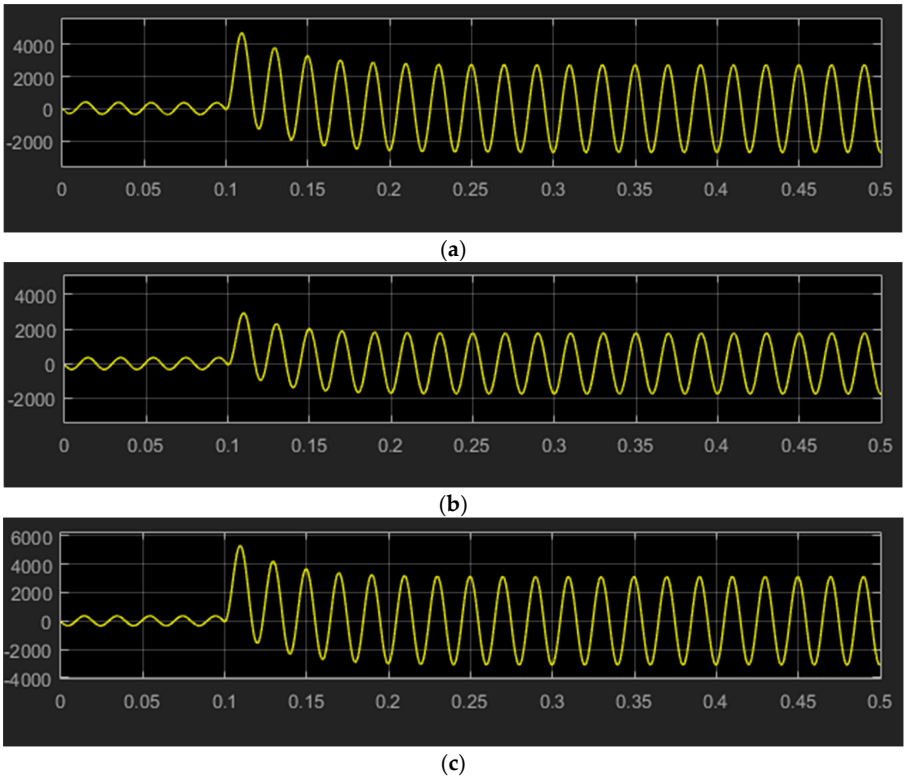
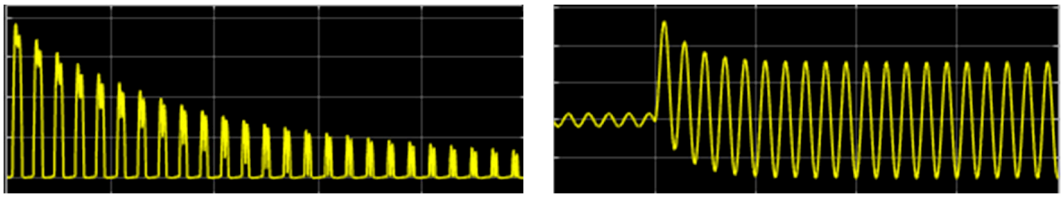


Figure 8. Waveform of fault current: (**a**) Short circuit between turns; (**b**) Single phase-to-ground fault; (**c**) Three-phase short circuit.

Table 2. Waveform information of fault current.

Fault Type	Peak Value/mA
Short circuit between turns	4723
Single phase-to-ground fault	2919
Three-phase short circuit	5291



(**a**) (**b**)

Figure 9. Waveform comparison: (**a**) Inrush current; (**b**) Fault current.

3.4. Harmonic Analysis

FFT analysis can convert a signal into individual spectral components, and thereby provide frequency information about the signal. Therefore, it can be used for analysis of current waveforms. In this section, FFT analysis was used to analyze every order of the harmonic content of the inrush current and short current.

Total harmonic distortion (THD) is one way to gauge power supply quality. It indicates how much of a harmonic component the voltage and current waveforms contain. In Figure 10 and Table 3, the THD of fault current did not exceed 20%, while that of the inrush current was above 40% and even reached 80%. The information demonstrates that there is a large difference between the inrush current and fault current on the fundamental harmonic and THD. In addition, the same phase angle and residual fluxes are used in the inrush current samples, respectively. The data in Table 3 show that phase angle and residual flux can affect the fundamental harmonic and THD of the inrush current.

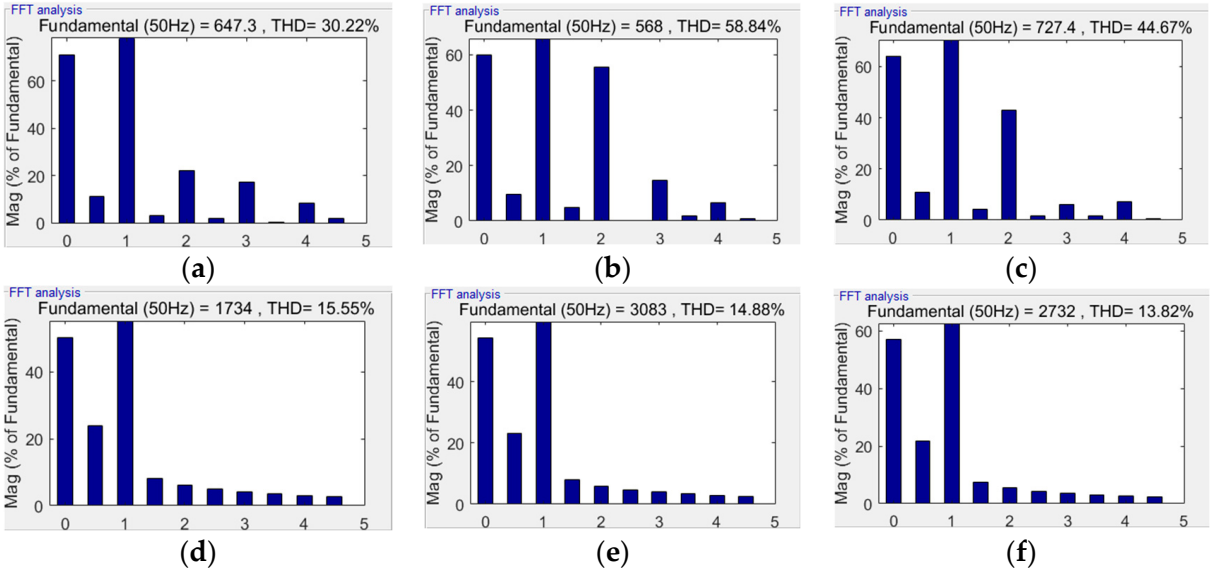


Figure 10. Harmonic analysis on current samples: (**a**) Inrush current under 0° phase angle and 0.5 pu residual flux; (**b**) Inrush current under 45° phase angle and 0.5 pu residual flux; (**c**) Inrush current under 45° phase angle and 0.8 pu residual flux; (**d**) Single phase-to-ground fault; (**e**) Three-phase short circuit; (**f**) Short circuit between turns.

Table 3. Analysis data of current samples.

	Fundamental	THD
Inrush current 1	384.7	44.11%
Inrush current 2	324.8	83.65%
Inrush current 3	661.2	50.60%
Fault current 1	1734	15.55%
Fault current 2	3083	14.88%
Fault current 3	2732	13.82%

3.5. Input of the Network

In the simulation of the current, various characteristics, such as dead angel in Figure 9, can be used to distinguish between the inrush and internal fault currents. However, extracting all features when modelling the network is not necessary. Too many eigenvalues being input will cause a slow converge or even failure of converge. The input should have features with strong regularity and high sensitivity to establish a network easily. Meanwhile, the computational efficiency and timeliness of the features must also be considered.

Figure 10 and Table 3 have shown that numerous high-order harmonics are generated when an inrush current occurs in a transformer, particularly the second harmonic. In contrast, the proportion of high-order harmonics in the fault current is small. Therefore, the percentage of harmonics can be selected as the input.

The proportion of high-order harmonics in the inrush current and fault current can be obtained through FFT analysis. The FFT analysis is shown in Figure 11. The inrush current contains a large number of high-order harmonics. However, the harmonic content above the fifth order is exceedingly low and can be ignored. Therefore, only the DC component and first to fifth harmonics were selected as inputs. The percentage of the first harmonic is always 100%; therefore, this item does not have any effect on identification. Finally, the input of the GA-BP network was determined as follows:

$$H = [H_0, H_2, H_3, H_4, H_5] \tag{23}$$

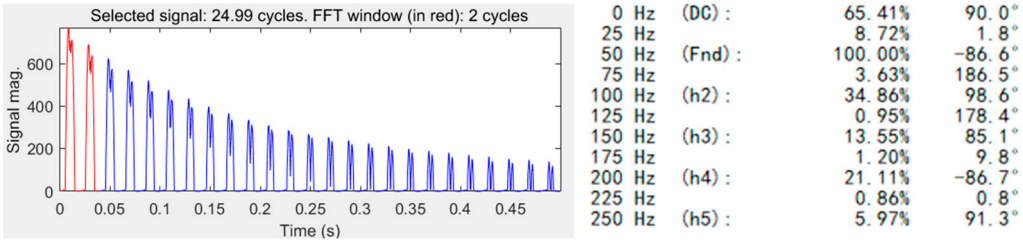


Figure 11. The proportion of harmonics in current presented by FFT analysis.

H_0 is the DC component, its range is $0 < H_0 < 2$, and H_2–H_5 are the second to fifth harmonic components with a range of 0–1. The number of output neurons in a neural network is related to its function of the neural network. To classify the inrush and fault currents, the output contained one node. When the input sample was an inrush current, the output was 1; otherwise, it was 0.

Different inrush currents could be obtained by changing the initial flux in the transformer and the initial phase of the power supply. Different fault currents could be obtained by changing the parameters of the circuit elements. Thus, a total of 200 samples was obtained. A portion of the sample is presented in Table 4.

Table 4. Part of training samples.

Current Type		H_0	H_2	H_3	H_4	H_5	Output
No-load energized	0°	0.654	0.349	0.136	0.211	0.06	1
	30°	0.582	0.626	0.256	0.054	0.047	1
	60°	0.54	0.789	0.511	0.243	0.061	1
	90°	0.145	0.737	0.411	0.869	0.413	1
	120°	0.533	0.795	0.538	0.281	0.085	1
	150°	0.571	0.65	0.282	0.04	0.068	1
	180°	0.656	0.345	0.137	0.208	0.055	1
Internal fault	Single-phase short circuit	0.503	0.062	0.041	0.031	0.025	0
	Two-phase short circuit	0.328	0.037	0.026	0.019	0.015	0
	Three-phase short circuit	0.543	0.059	0.039	0.029	0.024	0
	Turn-to-turn short circuit	0.509	0.049	0.033	0.025	0.019	0
Bus 8 fault	Fault between A phase and B phase	0.429	0.223	0.147	0.111	0.088	0
	Fault between A phase and C phase	0.254	0.181	0.119	0.089	0.071	0
	Ground fault	0.387	0.223	0.147	0.110	0.088	0
	Turn-to-turn short circuit	0.038	0.116	0.083	0.042	0.034	0
External fault	Single-phase short circuit	0.077	0.009	0.006	0.004	0.004	0
	Two-phase short circuit	0.421	0.022	0.021	0.014	0.013	0
	Three-phase short circuit	0.221	0.023	0.016	0.012	0.009	0
Inrush current occurred in Bus1 when fault occurred in Bus6	Fault between A phase and B phase	0.462	0.182	0.126	0.094	0.075	0
	Fault between A phase and C phase	0.197	0.147	0.096	0.071	0.057	0
	Ground fault	0.354	0.188	0.123	0.092	0.073	0

The case studies used to validate the proposal is taken from the FFT analysis. The input of network is $[H_0, H_2, H_3, H_4, H_5]$, which means 1st~5th order harmonic content of current. Output 1 presents means that the current is regarded as an inrush current and Output 0 is regarded as fault current.

In order to make the experimental results better reflect the performance of GSA-BP in practical applications, 100 current samples were taken from IEEE 14 bus system case (Figure 12) [37]. The simulating system represents an approximation of the American Electric Power system. It can simulate the possible fault conditions in real situations.

3.6. Samples Classification through the BP and GSA-BP Networks

The samples obtained using the Simulink simulation functioned as the input for the neural network. The maximum number of network cycles was 1000, and the target error was 0.0001 [38]. The structure of the BP network and transfer function is listed in Table 5.

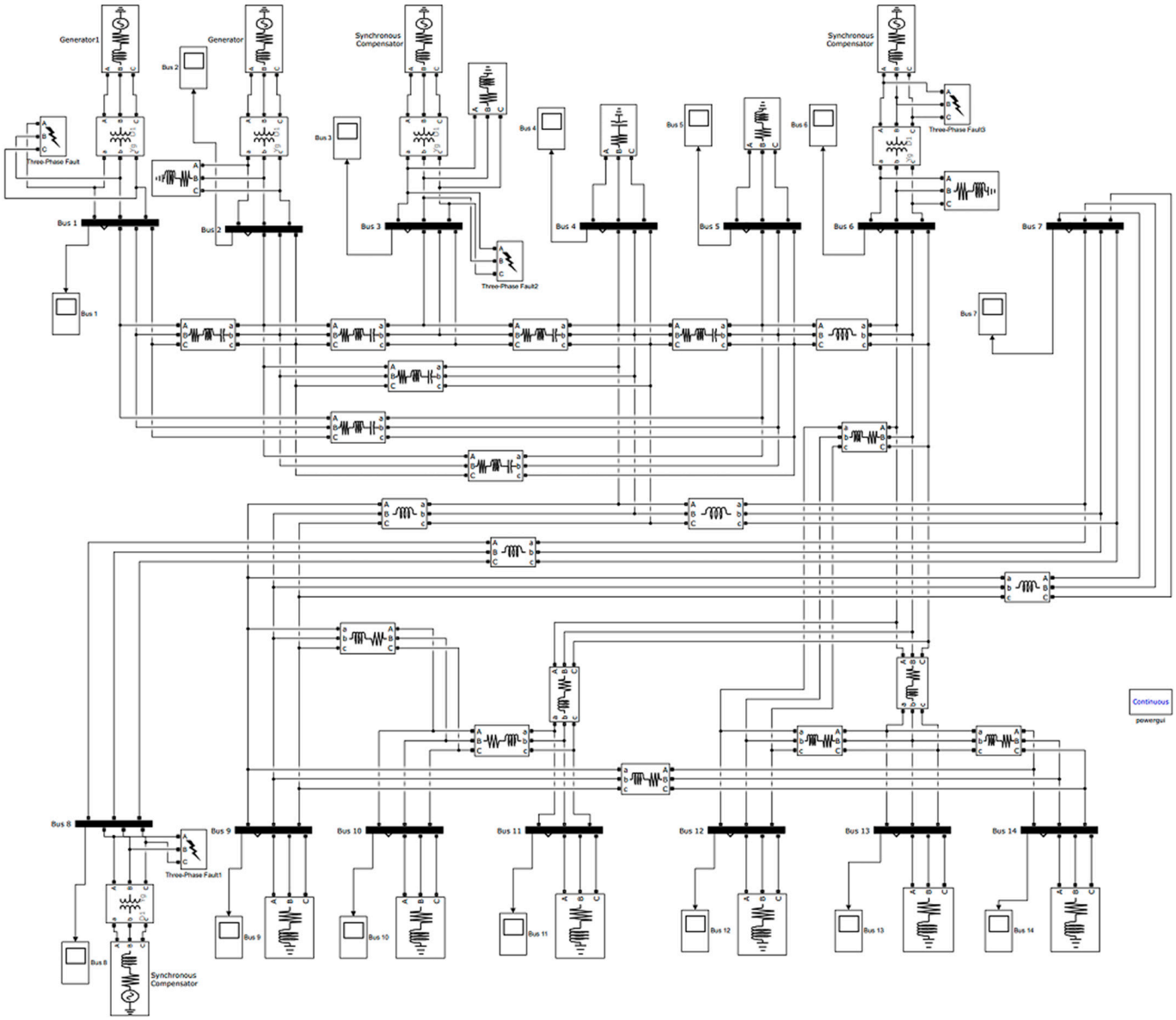


Figure 12. IEEE 14 bus system case.

Table 5. Structure and function of BP neural network.

BP Neural Network	Parameters
Number of input layer nodes	5
Number of hidden layer nodes	11
Hidden layer transfer function	Sigmoid function
Output layer transfer function	Linear function
Expected error	0.00001
Training function	Levenberg–Marquardt

The training results and test outputs are shown in Figure 13. The BP network reached convergence at 110 epoches.

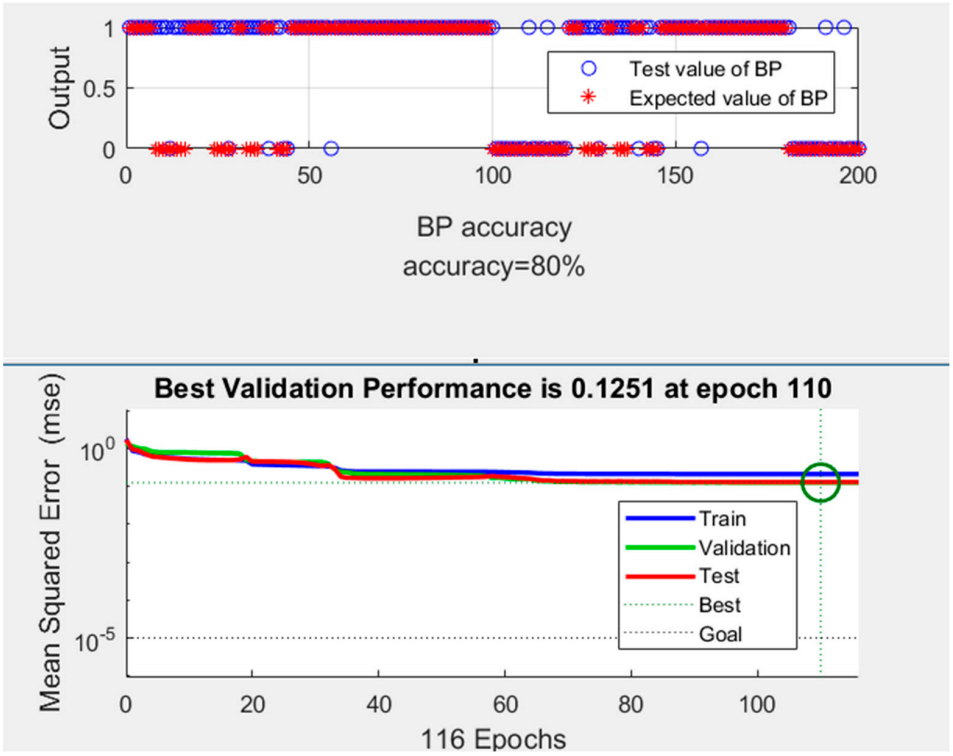


Figure 13. Testing results and test outputs of the BP network.

The fitness curve and training results of GSA-BP are shown in Figures 14 and 15. According to the results, the GSA-BP network converged at 20 epoches.

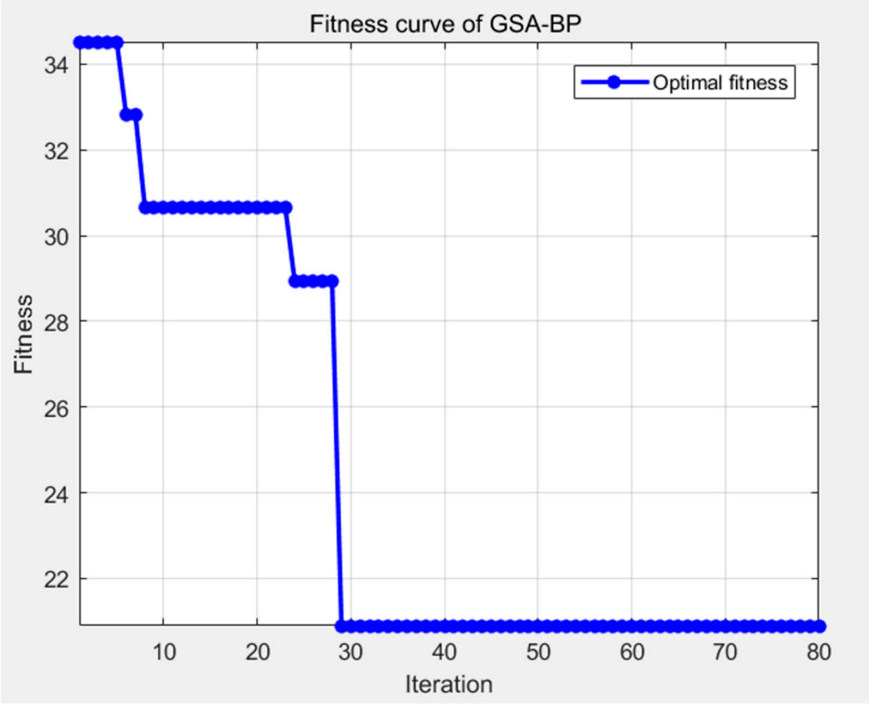


Figure 14. Fitness curve of the GSA-BP.

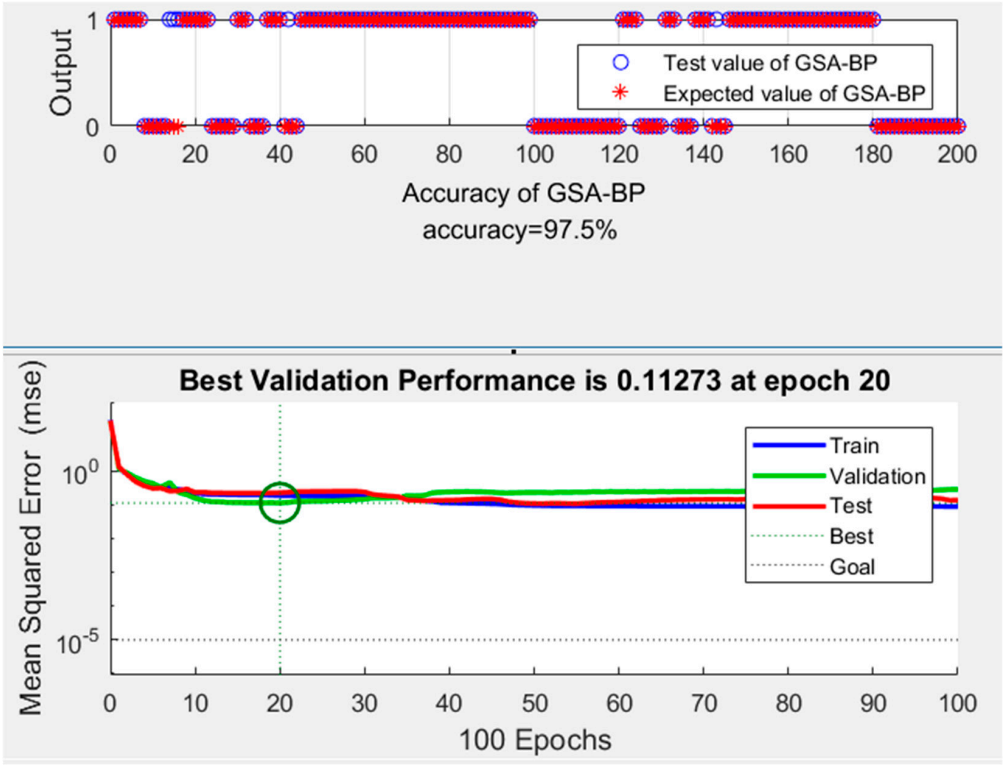


Figure 15. Testing results and test outputs of GSA-BP network.

The same test samples were input into the GSA-BP network. The outputs are shown in Figure 15.

4. Comparative Study

The current simulation was executed using a simulating model in MATLAB 2017a. The current simulation involved various possible fault currents, including internal faults and no-load, energized on internal faults in different scenarios to bring it closer to reality.

Actually, the BP neural network has been applied in transformer fault diagnosis. However, the accuracy and diagnosis speed of the basic BP network are not enough because the optimal state of the mis-operation does not occur. It is necessary to combine different algorithms to deal with the application of different scenarios. In this study, the proposed algorithm was based on a BP network. A normal BP algorithm is not sufficient for the current examination because the experimental results indicated that its recognition ability is unreliable. GA and SA algorithms were used to optimise the BP network for a more accurate and stable recognition.

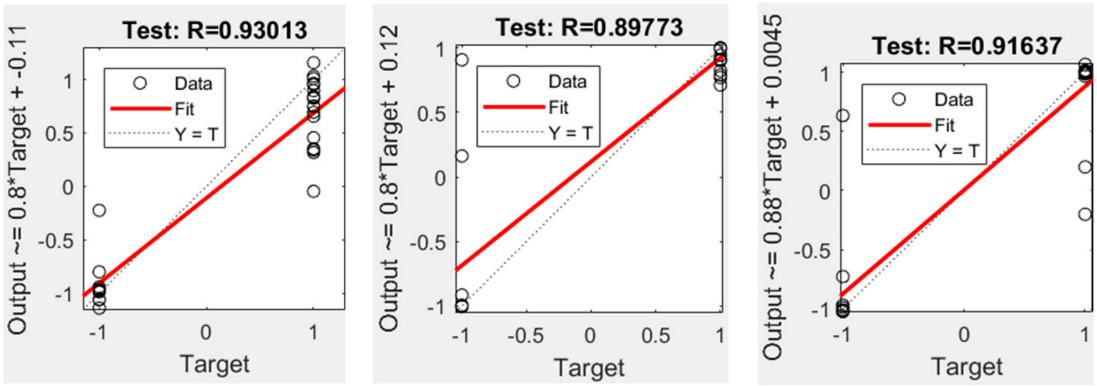
The comparison of waveforms and the statistics of harmonics show a sharp difference in the high harmonic content between inrush and fault currents. Therefore, the fundamental and 2nd–5th harmonics contents of current are used as characteristic values to input network. Further, the current generated in the simulation under no-load energized transformer fault in a different situation, and the no-load energized when the transformer fault occurred. All the samples, 200 in total, are classified by the optimized neural network and original network. This section shows the recognition speed and accuracy of the network. In addition, BP-related algorithms that have been proposed by other researcher are also presented for comparison in order to highlight the advantages of the proposed algorithm.

After the simulation, the performances of the BP and GSA-BP networks can be clearly observed. A comparison is presented in Table 6.

Table 6. Comparison of BP and GSA-BP networks.

	BP Network	GA-BP Network	GSA-BP Network
Number of epoches required to reach convergence	85	27	15
Current classification accuracy	77.5%	92%	97.5%

Table 6 shows that the current classification models established by the BP networks had identification accuracies of 80% for all samples. The accuracy of the GSA-BP network identification model was 97.5%. All training samples and test samples are taken from different parameters and scenarios, and the eigenvalues of each sample are different, so GSA-BP does not have overfitting in the test. In order to better reflect the ability of GSA-BP to respond to different situations, the regression curves under different conditions are listed (Figure 16).



(**a**) (**b**) (**c**)

Figure 16. Regression of test on different types of samples: (**a**) Inrush current; (**b**) Fault current; (**c**) Inrush current and fault current happen at the same time.

GSA-BP maintains a high degree of fitting in the testing of different data sets, which indicates that GSA-BP has a sufficiently stable performance in response to changes. In terms of accuracy, GSA-BP has an overwhelming advantage. In terms of convergence speed, GSA-BP also requires the least number of steps to reach convergence. The GSA-BP network achieved excellent results regardless of the convergence speed or accuracy of identifying the current. The simulation also showed that the algorithm can reduce the search range at a faster rate. The algorithm also provides a high local search efficiency. The classification model achieved an excellent performance improvement after the optimisation of these capabilities.

While BP-related identification is compared with the GSA-BP network, the algorithm proposed in the related articles is also listed in Table 7 for comparison on the accuracy and recognition speed [39,40].

These are similar current identification methods that have been proposed, and the table shows their accuracy in identifying an inrush current. Although they all have their own strengths in different fields, compared to GSA-BP, they all have lower accuracy. While ensuring high accuracy, GSA-BP takes only 0.186 s to identify a current sample on average. Compared with BP-related recognition methods, such as BP-Adaboost PNN, the recognition time is reduced by 34.5%. Additionally, the accuracy increased by 9.5%. It is unique to achieve such a recognition speed at the same level of accuracy. A comparative analysis indicated that the current classification based on BP and other proposed methods can discriminate between the inrush and internal fault currents. However, the high training speed and discrimination precision of GSA-BP are not available in other classification

methods. The GSA-BP network has broad application prospects for the classification of inrush currents.

Table 7. Comparison with different classification method proposed by related paper.

	PNN	BP-Adaboost	BP-Adaboost PNN	GA-BP	GSA-BP
Current classification accuracy	72%	81.5%	89.5%	87.5%	97.5%
Improvement achieved by GSA-BP network	25.5%	16%	8%	10%	0%
The average time required to identify a sample(s)	0.003	0.038	0.284	28.94	0.186

5. Conclusions

In this paper, a BP neural network was enhanced through optimisation based on simulated annealing and genetic algorithm, which improves the learning efficiency and global optimisation. GA has excellent global search capabilities, whereas the annealing algorithm has the ability to avoid premature convergence of the genetic algorithm. They significantly improve the structural performance and computing speed of the original network. The inrush and fault currents are simulated with different initial phase angles and remanence. Comparing the current waveforms shows that the initial phase angles and remanence can influence the peak and harmonic distortion rate of the inrush current. In addition, the current waveforms of the inrush current have intermittent angles, while the waveforms of the fault current are relatively smooth. The harmonic contents of the inrush and fault currents were used as the inputs to the neural network focusing on the second to fifth harmonics. Both inrush and fault currents were generated to provide experimental samples for the identification of neural networks. Compared with BP and GA-BP networks, the GSA-BP network has a higher iteration speed and improved accuracy with the recognition rate for all current samples is achieved at 97.5%. The application of the method requires the adjustment to the number of layers and neurons to be used.

Author Contributions: Conceptualization, Z.R. and N.N.B.M.; methodology, Z.R.; software, Z.R.; validation, Z.R. and N.N.B.M.; data curation, Z.R.; writing—original draft preparation, Z.R.; writing review and editing, Z.R., N.N.B.M. and H.A.I. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Acknowledgments: The authors of the article appreciate the referees for their valuable suggestions, which contributes to improving the paper.

Conflicts of Interest: The authors declare no conflict of interest.

List of Symbols and Abbreviations

BP	Backpropagation
GA	Genetic algorithm
FFT	Fast Fourier transform
I_{d2}	Second harmonic current in differential current
I_{d1}	Fundamental current in differential current
K_2	Second harmonic braking factor
MCDM	Multi-criteria decision-making
a	Output of neuron

w	Weights
p	Input of neuron
b	Deviation
t	Expected output
η	Learning rate
K	Length of the code
i_n	Input layer neuron
h_k	Hidden layer neuron
o_m	Output layer neuron
X_n	Input sample
Y_m	Network output
ω	Weights
θ	Threshold
E	Total error in the neural network
T_j^k	Ideal output
Y_j^k	Real output
K	Number of samples
P_{s_i}	Probability of each chromosome being selected
f_i	Fitness function value of the first chromosome
x	Individual
N	Population size
P_c	Crossover probability
$P_{m:}$	Mutation probability
f_{max}	Maximum fitness value in the current population
f_{min}	Minimum fitness value in the current population
f_v	Average fitness value of the current population
f'	Fitness value of the chromosome with larger fitness between two chromosomes participating in the crossover
P_c	Crossover probability
P_m	Mutation probability
$f_{\max}$	Maximum value of the fitness value in the current population
$f_{\min}$	Minimum value of the fitness value in the current population
f_v	Average fitness value of the current population
f'	Fitness value of the chromosome with larger fitness between two chromosomes participating in the crossover

Appendix A

Table A1. Introduction to electrical components.

Power Components	Function	Parameters
Three phase sourse	Provide continuous and stable sinusoidal AC voltage	Voltage, phase angle, resistance, inductance, etc.
Three-phase sourse RLC load	Simulate the parallel load	Load rated voltage, frequency, active power, reactive power, etc.
Three-phase VI measurement	Detect the voltage and current across the transformer	None
Three-phase breaker	Cut off the faulty circuit	Cutting off initial time, circuit breaker parameters, etc.
Three-phase transformer	Simulate transformer	Iron core, winding connection method, transformer capacity, frequency, etc.
Three-phase Fault	Simulate failure	fault, fault resistance parameters, etc.
Scope	Show the results	None

References

1. Kabbara, W.; Bensetti, M.; Phulpin, T.; Caillierez, A.; Loudot, S.; Sadarnac, D. A Control Strategy to Avoid Drop and Inrush Currents during Transient Phases in a Multi-Transmitters DIPT System. *Energies* **2022**, *15*, 2911. [CrossRef]
2. Habyarimana, M.; Dorrell, D.G.; Musumpuka, R. Reduction of Starting Current in Large Induction Motors. *Energies* **2022**, *15*, 3848. [CrossRef]
3. Abdusalam, O.; Ibrahim, A.; Anayi, F.; Packianather, M. New Hybrid Machine Learning Method for Detecting Faults in Three-Phase Power Transformers. *Energies* **2022**, *15*, 3905. [CrossRef]
4. Gunda, S.K.; Dhanikonda, V.S.S.S.S. Discrimination of Transformer Inrush Currents and Internal Fault Currents Using Extended Kalman Filter Algorithm (EKF). *Energies* **2021**, *14*, 6020. [CrossRef]
5. Marvasti, F.D.; Mirzaei, A. A Novel Method of Combined DC and Harmonic Overcurrent Protection for Rectifier Converters of Monopolar HVDC Systems. *IEEE Trans. Power Deliv.* **2018**, *33*, 892–900. [CrossRef]
6. Lee, B.; Lee, J.; Won, S.; Lee, B.; Crossley, P.; Kang, Y. Saturation Detection-Based Blocking Scheme for Transformer Differential Protection. *Energies* **2014**, *7*, 4571–4587. [CrossRef]
7. Cao, W.; Yin, X.; Zhang, Z.; Pan, Y.; Wang, Y.; Yin, X. Characteristic analysis of zero-mode inrush current of high-impedance transformer. *Int. J. Electr. Power Energy Syst.* **2020**, *117*, 105716. [CrossRef]
8. Okilly, A.H.; Kim, N.; Baek, J. Inrush Current Control of High Power Density DC–DC Converter. *Energies* **2020**, *13*, 4301. [CrossRef]
9. Dashti, H.; Davarpanah, M.; Sanaye-Pasand, M.; Lesani, H. Discriminating transformer large inrush currents from fault currents. *Int. J. Electr. Power Energy Syst.* **2016**, *75*, 74–82. [CrossRef]
10. Zhang, C.; Lu, Y. Study on artificial intelligence: The state of the art and future prospects. *J. Ind. Inf. Integr.* **2021**, *23*, 100224. [CrossRef]
11. Yazdani-Asrami, M.; Taghipour-Gorjikolaie, M.; Razavi, S.M.; Gholamian, S.A. A novel intelligent protection system for power transformers considering possible electrical faults, inrush current, CT saturation and over-excitation. *Int. J. Electr. Power Energy Syst.* **2015**, *64*, 1129–1140. [CrossRef]
12. Rahmati, A.; Sanaye-Pasand, M. Protection of power transformer using multi criteria decision-making. *Int. J. Electr. Power Energy Syst.* **2015**, *68*, 294–303. [CrossRef]
13. Lee, B.; Park, J.-W.; Crossley, P.; Kang, Y. Induced Voltages Ratio-Based Algorithm for Fault Detection, and Faulted Phase and Winding Identification of a Three-Winding Power Transformer. *Energies* **2014**, *7*, 6031–6049. [CrossRef]
14. Zhang, A.Q.; Ji, T.Y.; Li, M.S.; Wu, Q.H.; Zhang, L.L. An Identification Method Based on Mathematical Morphology for Sympathetic Inrush. *IEEE Trans. Power Deliv.* **2018**, *33*, 12–21. [CrossRef]
15. Marques, J.P.; Lazaro, C.; Morais, A.P.; Cardoso, G. A reliable setting-free technique for power transformer protection based on waveform transform. *Electr. Power Syst. Res.* **2018**, *162*, 161–168. [CrossRef]
16. Zhang, L.L.; Wu, Q.H.; Ji, T.Y.; Zhang, A.Q. Identification of inrush currents in power transformers based on higher-order statistics. *Electr. Power Syst. Res.* **2017**, *146*, 161–169. [CrossRef]
17. Deng, Y.; Lin, S.; Fu, L.; Liao, K.; Liu, L.; He, Z.; Gao, S.; Liu, Y. New Criterion of Converter Transformer Differential Protection Based on Wavelet Energy Entropy. *IEEE Trans. Power Deliv.* **2019**, *34*, 980–990. [CrossRef]
18. Almaghrabi, S.; Rana, M.; Hamilton, M.; Rahaman, M.S. Solar power time series forecasting utilising wavelet coefficients. *Neurocomputing* **2022**, *508*, 182–207. [CrossRef]
19. Huang, J.; Ling, L.; Xiao, Q. Research on weak signal detection method for power system fault based on improved wavelet threshold. *Energy Rep.* **2022**, *8*, 290–296. [CrossRef]
20. Narayanamoorthy, S.; Parthasarathy, T.N.; Pragathi, S.; Shanmugam, P.; Baleanu, D.; Ahmadian, A.; Kang, D. The novel augmented Fermatean MCDM perspectives for identifying the optimal renewable energy power plant location. *Sustain. Energy Technol. Assess.* **2022**, *53*, 102488. [CrossRef]
21. Valipour, M. Optimization of neural networks for precipitation analysis in a humid region to detect drought and wet year alarms. *Meteorol. Appl.* **2016**, *23*, 91–100. [CrossRef]
22. Dopazo, D.A.; Pelayo, V.M.; Fuster, G.G. An automatic methodology for the quality enhancement of requirements using genetic algorithms. *Inf. Softw. Technol.* **2021**, *140*, 106696. [CrossRef]
23. Sang, B. Application of genetic algorithm and BP neural network in supply chain finance under information sharing. *J. Comput. Appl. Math.* **2021**, *384*, 113170. [CrossRef]
24. Wang, L.; Bi, X. Risk assessment of knowledge fusion in an innovation ecosystem based on a GA-BP neural network. *Cogn. Syst. Res.* **2021**, *66*, 201–210. [CrossRef]
25. Ajmal, M.S.; Iqbal, Z.; Khan, F.Z.; Ahmad, M.; Ahmad, I.; Gupta, B.B. Hybrid ant genetic algorithm for efficient task scheduling in cloud data centers. *Comput. Electr. Eng.* **2021**, *95*, 107419. [CrossRef]
26. Aygun, H.; Turan, O. Application of genetic algorithm in exergy and sustainability: A case of aero-gas turbine engine at cruise phase. *Energy* **2022**, *238*, 121644. [CrossRef]
27. Zou, M.; Xue, L.; Gai, H.; Dang, Z.; Wang, S.; Xu, P. Identification of the shear parameters for lunar regolith based on a GA-BP neural network. *J. Terramech.* **2020**, *89*, 21–29. [CrossRef]
28. Esnaashari, M.; Damia, A.H. Automation of software test data generation using genetic algorithm and reinforcement learning. *Expert Syst. Appl.* **2021**, *183*, 115446. [CrossRef]

29. Wang, C.; Guo, C.; Zuo, X. Solving multi-depot electric vehicle scheduling problem by column generation and genetic algorithm. *Appl. Soft Comput.* **2021**, *112*, 107774. [CrossRef]
30. Xiao, H.; Tian, Y. Prediction of mine coal layer spontaneous combustion danger based on genetic algorithm and BP neural networks. *Procedia Eng.* **2011**, *26*, 139–146. [CrossRef]
31. Muñoz, A.; Rubio, F. Evaluating genetic algorithms through the approximability hierarchy. *J. Comput. Sci.* **2021**, *53*, 101388. [CrossRef]
32. Shyla, M.K.; Kumar, K.B.S.; Das, R.K. Image steganography using genetic algorithm for cover image selection and embedding. *Soft Comput. Lett.* **2021**, *3*, 100021. [CrossRef]
33. Yu, F.; Xu, X. A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. *Appl. Energy* **2014**, *134*, 102–113. [CrossRef]
34. Nachaoui, M.; Afraites, L.; Laghrib, A. A Regularization by Denoising super-resolution method based on genetic algorithms. *Signal Process. Image Commun.* **2021**, *99*, 116505. [CrossRef]
35. Oyanagi, R.; Noda, T.; Ichikawa, M. A Method for Estimating the Current-Flux Curve of a Single-Phase Transformer for Electromagnetic Transient Simulations of Inrush Currents. *Electr. Eng. Jpn.* **2018**, *204*, 13–24. [CrossRef]
36. Jung, B.I.; Cho, Y.S.; Park, H.M.; Chung, D.C.; Choi, H.S. Comparison of the quench and fault current limiting characteristics of the flux-coupling type SFCL with single and three-phase transformer. *Phys. C Supercond.* **2013**, *484*, 258–262. [CrossRef]
37. Illinois Center for a Smarter Electric Grid. IEEE 14-Bus System. August 1993. Available online: https://icseg.iti.illinois.edu/ieee-14-bus-system/ (accessed on 18 January 2023).
38. Liang, W.; Wang, G.; Ning, X.; Zhang, J.; Li, Y.; Jiang, C.; Zhang, N. Application of BP neural network to the prediction of coal ash melting characteristic temperature. *Fuel* **2020**, *260*, 116324. [CrossRef]
39. Bagheri, S.; Moravej, Z.; Gharehpetian, G.B. Classification and Discrimination Among Winding Mechanical Defects, Internal and External Electrical Faults, and Inrush Current of Transformer. *IEEE Trans. Ind. Inform.* **2018**, *14*, 484–493. [CrossRef]
40. Yan, C.; Li, M.; Liu, W. Transformer Fault Diagnosis Based on BP-Adaboost and PNN Series Connection. *Math. Probl. Eng.* **2019**, *2019*, 1019845. [CrossRef]