IOP Conf. Series: Materials Science and Engineering **439** (2018) 032120 doi:10.1088/1757-899X/439/3/032120

Application of ACO-LMBP Hybrid Neural Network Algorithm in Image Denoising

Hai jun Wang, Jin Tao, Menke Neimule

Ordos Institute Of Technology, Inner Mongolia Ordos,017000,China

Abstract. In order to overcome the disadvantages of poor global search ability, slow convergence speed and easy to fall into local minimum in the traditional BP neural network in image denoising, a hybrid ACO-LMBP neural network image denoising algorithm based on ant colony algorithm and LMBP algorithm is proposed. ACO-LMBP hybrid neural network algorithm has both the high speed of LMBP algorithm and the global nature of ACO algorithm. It can improve the problems of BP algorithm model very well. By comparing with the image denoising effect of Wiener filtering, BP, LMBP and PSO-LMBP model, the denoising model using the ACO-LMBP neural network algorithm has better denoising effect.

1. Introduction

Image denoising problem is one of the classic problems in image processing research. The research of denoising method is receiving more and more scholars' attention. Most of the traditional image non blind denoising methods are inverse solutions through point diffusion functions, such as median filtering, Wiener filtering, mean filtering and so on. The denoising effect of these methods is good, but the point spread function is difficult to solve, so the practical application scope of these methods is limited. The opposite is the image blind denoising method, which can estimate the effective point spread function without any prior information and then solve the problem. This type of blind denoising method is widely used at present, but it also has problem whicn are a large amount of computation, local convergence, and the solution is not unique [1]. Therefore, the traditional image denoising method can't meet the needs, many scholars began to explore new and more effective methods, especially the neural network method for image denoising is the most active. The neural network has the ability of large-scale parallel processing, and has good fault tolerance and associative memory ability. In recent years, it has been widely used. It is applied to research fields such as image feature extraction, image denoising, image segmentation and image recognition.

Back propagation (BP) algorithm is the most widely used neural network algorithm in the field of image processing. Its operation is based on the principle of gradient descent, so there are many problems such as weak global search ability, slow convergence speed and easy to fall into local minimum. In order to solve these problems, a Levenberg Marquardt (LM) BP algorithm is often used to model the algorithm. This algorithm combines the local convergence of the Gauss-Newton method and the global convergence of the gradient descent method. This method is very effective for image processing efficiency. In the BP algorithm [2,3], but because it can not fundamentally solve the BP algorithm initial weight threshold randomness, it is easy to fall into the local minimum value problem, so the model optimization effect is limited. Ant colony optimization (ACO) is a population-based heuristic algorithm proposed by Italian scholar M. Dorigo et al. in 1991. A large number of studies have shown that using this algorithm to optimize the initial weight and threshold of BP neural network can effectively improve the efficiency of BP neural network operation[4]. Based on this, a hybrid

ACO-LMBP algorithm combining ant colony algorithm and LMBP algorithm is proposed and applied to image denoising. Experimental results show that compared with Wiener filtering, BP, LMBP and PSO-LMBP algorithm, ACO-LMBP algorithm can achieve better results in modeling image denoising.

2. The principle of ACO-LMBP algorithm

2.1 LMBP algorithm introduced

The essence of the BP neural network optimization process is to find the solution corresponding to the minimum error of the problem to be solved. The mathematical description is shown in equation (1):

$$\begin{cases} \min E(w,v,\theta,r)=\dfrac{1}{N}\sum_{k=1}^{N}(y_k-t_k)^2 \\ s.t. w\in R^{m\times p},\quad v\in R^{p\times n},\quad \theta\in R^{p},\quad r\in R^{n} \end{cases} \tag{1}$$

Among them, w,v,θ,r are the weight and threshold of the desired structure of the *m-p-n* neural network, y_k is the actual output of the kth sample, and the kth sample of the t_k is the expected output. The solution to the final problem is the optimal weight and threshold parameter combination of the model. The basic idea is: to reverse the output error $E(w,v,\theta,r)$ in a certain way through the hidden layer to the input layer, and distribute the error $E(w,v,\theta,r)$ to all the nodes of each layer, so as to obtain the error signal of each layer node, according to this signal corrects the weight and threshold w,v,θ,r of each node. The BP algorithm uses the gradient descent method to solve the weight and threshold w,v,θ,r ,this method runs faster in the initial stage, however, in the later stage of the run, the output error is close to the optimal value. At this time, the algorithm runs extremely slowly and if the error surface is a multidimensional space, surfaces may fall into a local minimum during the training process. The change from the point to the multi-direction will increase the training error, which will make the weight and threshold w,v,θ,r parameter training unable to escape the local minimum value, the model training convergence speed will be slower and the generalization ability will be worse[5].

The LMBP algorithm is different from the basic BP algorithm in the process of correcting weight and threshold using local sample information, the LMBP algorithm uses all the sample information to correct the combined solution $x(k)$ of the weight threshold, he calculation of new solution $x(k+1)$ is shown in formula (2), where $\Delta x=-(J^T(x)J(x)+\mu I)^{-1}J^T(x)e(x)$ is the weight threshold change. $J(x)$ is the Jacobian matrix, the calculation is shown in equation (3),the $e_i(x)$ in the formula is the error of the *i* node, and *I* is the unit matrix, μ （ $\mu>0$ ） is proportional coefficient.

$$x(k+1)=x(k)+\Delta x \tag{2}$$

$$J(x(k)=\begin{bmatrix} \dfrac{\partial e_1(x(k))}{\partial x_1(k)} & \cdots & \dfrac{\partial e_1(x(k))}{\partial x_n(k)} \\ \cdots & \cdots & \cdots \\ \dfrac{\partial e_n(x(k))}{\partial x_1(k)} & \cdots & \dfrac{\partial e_n(x(k))}{\partial x_n(k)} \end{bmatrix} \tag{3}$$

When $\mu\to 0$, $\Delta x=-(J^T(x)J(x))^{-1}J^T(x)e(x)$,at this time, the LM algorithm becomes the Gauss-Newton algorithm.When $\mu\to\infty$, $J^T(x)J(x)$ is far less than μI , negligible, $\Delta x=-(\mu I)^{-1}J^T(x)e(x)$, at this point the LM algorithm becomes a gradient descent method. Thus, it can be seen that by adjusting the ratio coefficient μ , while maintaining a fast descent speed, LM can adaptively adjust between Gauss-Newton method and gradient descent method, so that each iteration is no longer along a single negative gradient direction, so LMBP algorithm has high optimization efficiency[6]. Fig.1 and 2 are image denoising using gradient descent algorithm, additional momentum method (tradingdm), adaptive learning rate method (traingdx) and LM algorithm (trainlm) under the same training parameters. The training error curve and local amplification results of the model can be seen from the figure, the optimization efficiency of the LM algorithm is obviously better than other BP algorithms.

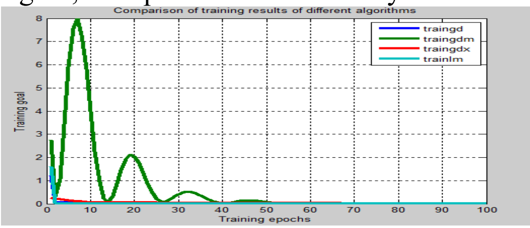

Fig.1. Comparison of training results of different BP algorithms

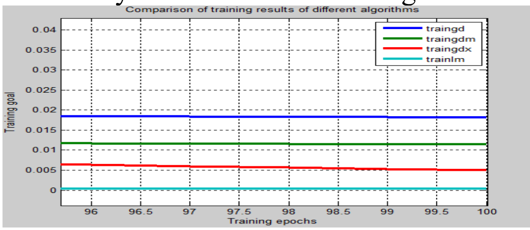

Fig.2. Local amplification of different training results

2.2 The core idea of ACO-LMBP algorithm

The core idea of the ACO-LMBP algorithm construction is to combine the weight threshold parameter w,v,θ,r of the BP model into a vector X as the initial parameter of the ant individual. Then the initial solution set of N ants is $I_X=\{X_1,X_2,\cdots,X_N\}$, the i ant colony starts from the initial position X_i,targeting the optimal fitness $F(X_{best})$, there are several paths between the initial location I_X and the target $F(X_{best})$ of the ant colony, that is, the combination of multiple w,v,θ,r. The ants release pheromones along the way to find the target. If there are a large number of ants passing through a path, the concentration of pheromone will be significantly higher than other paths, and more ants will be attracted to this path, ie positive feedback mechanism. As time goes on, the pheromone on the path will volatilize, and the concentration of pheromone will decrease, and the ant that chooses this path will decrease, ie, the negative feedback mechanism [7]. Through the constant interaction of positive and negative feedback mechanisms, the final ant colony will find an optimal path set X_i such that $F(X_i)=F(X_{best})$. At this moment, X_i is the optimal parameter combination of the optimal weight and threshold w,v,θ,r. After decoding it, the LMBP algorithm is used for the two optimization to generate the weight and threshold parameter of the final ACO-LMBP algorithm model.

2.3 ACO-LMBP algorithm step design

Step1 Parameter initialization----The structure of the BP model to be solved is m-p-n, In the [-1,1] interval, d is selected randomly as a solution of the initial weight and threshold w,v,θ,r of the BP model, in which d is the encoding length of the ownership threshold, and its formula is shown as formula (4). Let s be the number of ants. At t, the pheromone of element i in set I_X is $\tau_i(t)$, and $\Delta\tau_i(t)=0$. Let the current iteration number NC=0, the maximum number of iterations be NC_max , and set the learning rate and training error of the BP algorithm.

$$d=mp+pn+p+n \tag{4}$$

Step2 Fitness function determination----As the ant colony algorithm in the optimization process is the shortest path corresponding fitness as the optimal fitness, so according to the optimization objectives define the optimization problem of the function described as follows:

$$\begin{cases} F(w,v,\theta,r)=E(w,v,\theta,r)=\dfrac{1}{N}\sum\limits_{k=1}^{N}(y_k-y'_k)^2 \\ s.t. w\in R^{m\times p},\quad v\in R^{p\times n},\quad \theta\in R^p,\quad r\in R^n \end{cases} \tag{5}$$

Where y_k is the actual value of the k-th sample, y'_k is the output value of the k-th sample model, and N is the number of training samples. The ant colony algorithm is used to solve the above quadratic nonlinear optimization problem, that is, the optimal weight threshold w,v,θ,r parameter combination of the BP model is obtained.

Step3 Move traversal----The ants in the ant colony are independent of each other when looking for the shortest path. The transition probability that the *k*-th ant chooses to transfer to the *j*-th parameter point at time *t* is shown in formula (6), the roulette method is used to select the next parameter point. When the ant completes the selection for all parameter points, it goes to the final target.

$$P(\tau_j^k(t)) = \begin{cases} \dfrac{\tau_j(t)}{\sum_{u=1}^{d}\tau_u(t)}, 1 \le j \le d \\ 0 \quad , otherwise \end{cases} \tag{6}$$

Step4 Pheromone Update----As the number of iterations increases, the pheromone on each path that the ant passes through will also change. Here, only the amount of information on the shortest path that the ant passes through is updated. After updating, the NC is judged. Whether the maximum number of iterations *NC_max* is reached, if *NC==NC_max*, go to Step5; otherwise, go to Step3. The rules for updating the optimal path pheromone are shown in equation (7) [8].

$$\tau_i(t+n) = (1-\rho)\tau_i(t) + \Delta\tau_i(t+n) \tag{7}$$

In the formula:

$$\Delta\tau_i(t+n) = \sum_{k=1}^{d}\Delta\tau_i^k(t+n) \tag{8}$$

$$\Delta\tau_i^k(t+n) = Q/e^k \tag{9}$$

$$e^k = E(w,v,\theta,r) = \frac{1}{N}\sum_{i=1}^{N}(y_i^k - y_i)^2 \tag{10}$$

ρ is the volatile coefficient of pheromone, *d* is the number of parameters to be optimized, and *Q* is the adjustment speed coefficient of pheromone. y_i^k and y_i are the actual output and expected output of the *k*th ant, *N* is the number of training samples, and e^k is the fitness evaluation result of the kth ant.

Step 5 Weight Threshold Parameter Determination----Decode the value corresponding to the shortest path optimized by the ant colony algorithm to obtain the BP neural network initial weight and threshold w,v,θ,r , and assign it to the neural network image denoising model based on LMBP algorithm. Optimize to determine the final model parameters.

3.Sample structure and BP model parameters determination

3.1 Sample structure

For the image sample information, a sliding window with a block idea is used for traversal acquisition. The commonly used sliding windows are four squares (2×2), nine squares (3×3), sixteen squares (4×4).Fig.3-5 show the results of dividing the image area using the four square, nine square, and sixteen square . In view of the difficulty in selecting the center area and the starting position of the window is difficult to determine of the four squares and the sixteen squares, the nine squares not only separate the center of the image separately, but also the edge area of the noise image is easily determined, therefore, the input layer node is constructed by using the sliding window of nine squares.For the output sample, because the output of the model is a noiseless pixel in the center of the sliding window, the number of nodes of the output layer is 1, so that the model input and output sample of the <9 input, the 1 output > is constructed.


Fig.3 (a)


Fig.4 (b)

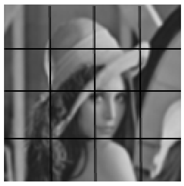
Fig.5 (c)

a.image division of four palaces b. the image division of the ninth Palace c. image division of sixteen palaces

3.2 Parameter determination of BP model

Neural network related theory has proved that as long as there are enough hidden layer nodes, BP neural network with single hidden layer can approximate any nonlinear function with finite discontinuity with arbitrary precision, so the input and output layers are connected. In the case where the point has been determined, the next step is to determine the hidden layer node. In this paper, the golden section method is used to determine the number of hidden layer nodes by using equation (11), where m and n represent the number of nodes in the input layer and the output layer, respectively. Calculate the error values of 0.618 and 0.382 points. After the golden section method, the number of hidden layer nodes is determined to be 12.

$$a = \frac{n+m}{2} \le r \le (n+m)+10 = b \tag{11}$$

Other related parameter settings: BP learning rate *lr*=0.05, training number *epochs*=500, training target *goal*=0.0001, training algorithm using trainlm; PSO algorithm population size *n*=30, learning factor *c*1=*c*2=1.5, speed *Vmax*= 1, *Vmin* = -1, number of iterations *Nmax* = 100; ACO algorithm ant colony size *ant* = 30, pheromone volatilization coefficient ρ = 0.8, transition probability constant P_0 = 0.2, number of iterations *NC_max* = 50.

4.Experimental analysis

The Gauss noise image of 150 * 150 Brain and Cameraman is used as the test sample. At the same time, Wiener filtering, BP, LMBP, PSO-LMBP, and the ACO-LMBP algorithm modeling proposed in this paper are used for denoising experiments. Fig.6-12 is the result of the correlation denoising experiment performed on the Brain diagram. It can be seen from the figure that Winner filtering does not eliminate noise well, and the image quality is not significantly improved than the noisy graph. The BP algorithm model not only fails to eliminate noise, but also makes the image details more fuzzy, because the LMBP, PSO-LMBP and ACO-LMBP have achieved good denoising effect. The same conclusions were obtained from the related experiments of the Cameraman diagram from Fig.13-19. In order to further quantitatively evaluate several algorithms, a structural similarity (SSIM) function and a peak signal-to-noise ratio (PSNR) function are used as evaluation criteria. Table 1 gives the evaluation results based on these two evaluation functions. It can be seen from the table that the ACO-LMBP denoising model proposed in this paper is the largest in terms of structural similarity and peak signal-to-noise ratio, and it is said that the denoising effect is the best.

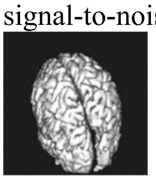
Fig.6 (a)

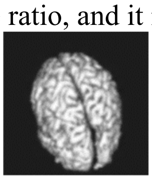
Fig.7 (b)

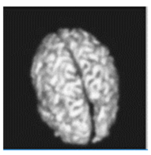
Fig.8 (c)

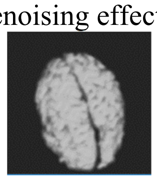
Fig.9 (d)

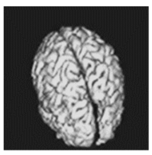
Fig.10 (e)

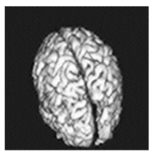
Fig.11 (f)

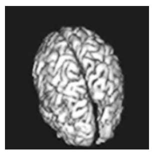
Fig.12 (g)

IOP Conf. Series: Materials Science and Engineering **439** (2018) 032120 doi:10.1088/1757-899X/439/3/032120

      

Fig.13 (a) Fig.14 (b) Fig.15 (c) Fig.16 (d) Fig.17 (e) Fig.18 (f)

Fig.19 (g)

a. original image b. noisy image c. Winner d. BP e. LMBP f. PSO-LMBP g. ACO-LMBP

Table 1 Comparison of Denoising Effects of Different Algorithms

Denoising algorithm	Brain image		Cameraman image	
	SSIM	PSNR	SSIM	PSNR
Wiener	0.8448	22.7727	0.8230	23.8284
BP	0.7102	19.6865	0.7499	22.4799
LMBP	0.9708	27.6582	0.9457	30.6658
PSO-LMBP	0.9763	29.6807	0.9597	31.1680
ACO-LMBP	0.9784	30.0721	0.9622	31.4940

5. Conclusion

Aiming at fuzzy image denoising, an ACO-LMBP image denoising algorithm based on improved BP algorithm LMBP algorithm and ant colony algorithm is proposed. The algorithm firstly uses the ACO algorithm to obtain the approximate optimal solution of BP algorithm, and then uses LMBP to perform two, suboptimal to determine the final algorithm model parameters. The ACO-LMBP algorithm can dynamically control the mutual conversion process between global search and local search when the parameters are solved, and has better convergence. The experimental results show that the image denoising model based on ACO-LMBP algorithm has certain feasibility, and the denoising effect is better than Wiener filtering, BP, LMBP and PSO-LMBP algorithm models. Therefore, it has in-depth research value, and at the same time, the model related parameters can be modified. The ACO-LMBP algorithm can also be applied to other image processing research fields, the algorithm has the value of promotion research.

Acknowledgment

This work has been supported by China National Natural Science Foundation Project (61741509), Scientific research project of Inner Mongolia higher education institutions (NJZZ07007, NJZY16382).

References

[1] Wang H, Cai N.(2006). Image restoration method based on neural network. Infrared & Laser Engineering, vol.10(35),pp:121-125

[2] Tripathi, P. (2014). Image compression enhancement using bipolar coding with lm algorithm in artificial neural network. Revista Latinoamericana De Estudios Educativos,vol.36,pp:159-181.

[3] Xu, Q., & He, J. (2011). Lmbp algorithm of remote sensing image classification improved by kernel functions. Journal of Image & Graphics,vol.16(12), pp:2206-2210.

[4] Lu H F&Gu X S.(2016).A Two Level Information Fusion Algorithm Based on Ant Colony Neural Network. Journal of Shanghai Jiaotong University,vol.50,pp:1323-1330

[5] MIAO Xinying，CHU Jinkui，DU Xiaowen.(2011).Application of LM-BP neural network in predicting dam deformation.Computer Engineering and Applications，vol.47(1),pp：220-222.

[6] Jing, L., Xie, R., Zhou, W., Liu, Y., Yin, S., & Zhang, C. (2017). Application of lm-bp neural network in simulation of shear wave velocity of shale formation. Journal of China University of Petroleum.vol.41(3),pp:75-81
[7] Zhu S x,Guo Z y.(2016).Study on Displacement Monitoring Model for Concrete Dams Based on Improved ACO-BP. Journal of China Three Gorges University(Natural Sciences),vol.38(5),pp:19-23
[8] Wang, S., Yan-Ming, L. I., Wen-Liao, D. U., Jin, Y., & Liu, C. L. (2012). State performance evaluation for the main pump bearing of pump truck based on ant colony optimization of neural network. Journal of Shanghai Jiaotong University.vol.46(4),pp:596-560