Indonesian Journal of Electrical Engineering and Computer Science
Vol. 20, No. 1, October 2020, pp. 132~137
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v20.i1.pp132-137

An optimise ELM by league championship algorithm based on food images

Salwa Khalid Abdulateef[1], Taj-Aldeen Naser Abdali[2], Mohanad Dawood Salman Alroomi[3]
Mohamed Aktham Ahmed Altaha[4]
[1,3,4] Computer science and mathematics college, Tikrit University, Iraq
[2]Centre for Cyber Security, Universiti Kebangsaan Malaysia, Malaysia

Article Info

Article history:

Received Feb 4, 2020
Revised Apr 6, 2020
Accepted Apr 20, 2020

Keywords:

Extreme learning machine
ELM
LCA algorthm
Food images
Optimazition

ABSTRACT

This paper presents an optimisation of extreme learning machine by league championship algorithm based on food images. Extreme Learning Machine (ELM) is an effective classifier because of the performance which is higher than other classifiers' aspects. However, some important drawbacks still work as a hindrance like failure of optimal selection weights for the weights of the input-hidden layer and the output of the threshold. In spite of the wide number of problem-solving attempts, there was no solution to be considered effective. This paper presents the approach of hybrid learning and the League Championship Algorithm is used by for the purpose of selecting the input weights and the thresholds outputs. The experimental outcomes showed that the performance of proposed technique is superior as compared according to different scenarios of the measures to benchmark. The proposed method has achieved an overall accuracy of 95% for UEC food 100 dataset and 94% for UEC food 256 dataset comparing with 94% and 80% for baseline approaches.

Corresponding Author:

Taj-Aldeen Naser Abdali,
Centre for Cyber Security,
Faculty of Information Science and Technology (FTSM),
Universiti Kebangsaan Malaysia,
43600 UKMBangi, Selangor, Malaysia.
Email: tag_naser2002@yahoo.com

1. INTRODUCTION

In ELM, the initiation of the hidden nodes was random and thereafter fixed without tuning repetitively and it has been introduced by [1-3], it is not required for the hidden nodes in ELM to be neuron identical. The connections between the output and the hidden layer are the only free parameters needed to be learnt. As follows, the formulation of the ELM appears as a linear-in-the-parameter model and it boils down to find out a solution for a linear system [4].

ELM has been used in the purpose of items' categorisation. It considered as a neural network of feed-forward single hidden layer and the generating of its input weights and threshold goes randomly. The least-square method calculates the output weights of ELM and it is a reason why the ELM shows a high speed on both testing and training [1, 5]. Simplicity and efficiency in terms of execution are the merits of this approach and [6], that is because of the training repetition and more ability to capture knowledge from training dataset that it has as one entity. In addition, the approach seems to perform much better with several kernel functions and non-linear activation, whereas some exception of SVM which used kernels only [7, 8].

The optimisation area of the ELM has been exposed by numerous researchers. The self-adjusting extreme learning machine, SA-ELM, has been presented and launched by some of the well-known researchers like [9]. The objective function values have been minimalized by the adjustment of the biasness of the hidden layer and the input-weights of ELM with "learning and teaching phase", and that is based on the conception of the improved teaching learning-based optimisation. The eight benchmarks functions applied the SA-ELM and test its feasibility and validity. Generalisation performance and regression accurateness appear better with SA-ELM compared to ELM and fast learning network. Though, its process applies more complexity and much running time than the process applies in ELM. Various optimisation-based ELM, in addition to SA-ELM, were found in the literature: the input weights and bias of ELM have been optimised by Sun [10], based on Particle Swarm Optimisation (PSO). Bat algorithm are used by both [11], for optimisation of the biases and weights of ELM. Competitive swarm optimizer has been applied by [12], to optimise the standards of the hidden neurons and the input weights of the Extreme Learning Machine. An active operator's particle swam optimisation algorithm have been used by [13], in order to obtain an optimal set of initial parameters for KELM, therefore, to create a classifier for the optimal KELM.

A memetic algorithm of ELM base has been proposed by [14]. The local search strategy is been embedded into the global optimisation framework and this is by the proposed algorithm in order to achieve optimal network parameters. New type of particle swarm heuristic algorithm has been proposed by [15]. Which is called Multitask Beetle Antennae Swarm Algorithm, and the structures of Beetle Antennae Search algorithm (BAS) and Artificial Bee Colony (ABS) algorithm is its own inspiration. The memetic algorithm approach is processed for the purpose of optimising the biases of Extreme Learning Machine and the input weights.

Furthermore, the using meta-heuristic searching concept for optimisation the ELM classifier was used with kernel type of ELM. [16-18], have proposed quantum-based behaved particle swarm optimisation weighted multiple kernel ELM. Several studies have addressed the random problem [9, 19]. The literature contains numerous approaches were used for optimising the weights and biases of ELM. However, the results of these solutions have shown several gaps such as a difficult structure of the algorithm, long training time, slow, and local minima. Consequently, this problem will be tried to solve it by using a League Championship Algorithm (LCA).

According to [20], extensive experiments on various benchmarks, the author indicated that the League Championship Algorithm shows encouraging performance which advocates supplementary practical applications and developments to be worthy for an investigation in the upcoming studies. The results also showed that when compared to Differential Evolutionary algorithm (DE), genetic algorithm (GA) and (PSO), LCA performs better considerably, its simple structure and it has a fast running. This led us to select LCA to be used as an optimisation algorithm for the weights and biases on ELM. Numerous researchers have applied LCA in many fields [21-24].

This paper presents a new combination method of ELM with LCA as so called L-ELM weights additional to hidden biases, also the Moore-Penrose (MP) generalized inversely in the purpose of calculating the output weights. The LCA centres an optimisation of the SLFN output weights and forcing the hidden biases and input weights within a practical range for improving the convergence performance of ELM. The organisation of this paper shown as follows: Section 2 presents some preliminaries of LCA and ELM also introduces the proposed method. Whereas Section 3, contains experiment results which is given to exhibit proposed algorithm effectiveness. Finally, Section 4 offers the concluding remarks.

2. RESEARCH METHOD

The following section firstly providing the dicription on how ELM principels are working particulary the selection of algorithm. Secondly giving the enough illustration about LCA dose work. Lastly, we are presenting the enhanced model of L-ELM which is showing the adjustment of the hidden layer bias and input weights of ELM uses a hybrid technique as so called LCA, and to improve performance of ELM.

2.1. Extreme learning machine (ELM)

Considering that, each vector of features is regarded as one row in the data and its corresponding item encoding. The data consists of N random distinct samples (x_j, t_j):
where: $x_j = (x_{j1}, x_{j2}, \dots, x_{jn})^T$ and $t_j = (t_{j1}, t_{j2}, \dots, t_{jm})^T$
There is a possibility to model an activation function $g(x)$ and $\tilde{N}$ hidden layer neurons with standard Single Hidden Layer Feed-forward Network (SLFN) as follows:

$$\sum_{i=1}^{\tilde{N}} \beta_i \, g(w_i x_j + b_i) = t_i \tag{1}$$

where: $j = 1, \dots N$; $w_i = (a_{i1}, a_{i2}, \dots, a_{in})^T$

w_i is the weight connecting of i^{th} hidden knob and input, β_i is the weight connecting i^{th} hidden node and the output and b_i is the threshold of i^{th} hidden neuron.

Huang et al. [1], proved the possibility of differentiation in the activation function when the needed number of the hidden layer neurons is lower than the data size or $\widetilde{N} < N$. The above described training algorithm of the ANN is introduced in three following points:

a) Assign arbitrarily random bias and weights as w_i, and b_i.
b) Calculate the hidden layer output matrix.
c) Calculate the output weights β by using the (MP) generalised inverse of the hidden layer output matrix.

2.2. League championship algorithm (LCA)

This method improved by [25], and lately suggested the based algorithm of stochastic population for a worldwide continuous optimisation that attempts to imitate an environment of a championship in which synthetic teams play for some weeks as an artificial league (iterations). In each week the schedule of the league is given, and a several members indicate sport teams play in pairs and the result of the game is determined in an expression of lose and win, the playing strength (fitness value) is given along with the anticipated team formation (solution) developed by each team. An artificial modelling matches the analysis of each team which places the needed in its formation (generating a new solution). As well as, the championship and the match of the afterwards week continuous as a number of seasons (stopping condition) until the optimal solution attainment. Figure 1 illustrates the LCA steps as follows.

```
Initialise the league size(L)and the number of seasons(S);t=1
Generate a league schedule;
Initialise team formations//generatea population of L solution
        //and determine the playing strengths (fitness value)
While t ≤ S.(L-1)efine win_size   // Window size
   t=t+1;
   For i=1 to L
      Devise a new formation for team i for the forthcoming match
      If the new formation is the fittest one
    End for
    If mod(t,L-1)=0
       Generate a league schedule;
    End if
End while
```

Figure 1. The Steps of LCA

2.3. An optimisation extreme learning machine (L-ELM)

The hidden biases and the input weights (linking the input to the hidden layer), in ELM, are elected arbitrarily and the output weights (linking the hidden to the output layer) are decided analytically by using the MP. The formula used to inverse a non-square matrix.

$$X^{\dagger} = (XX^{T})^{-1}\, X^{T} \quad (2)$$

where X^T transpose of matrix X, $X^{\dagger}$ MP of matrix X

Additional to the higher generalization performance, ELM is not only faster than the traditional learning algorithm but as well it avoids several difficulties encountered by the methods of learning, such as learning rate, stopping criteria, local minima and learning epochs. However, in many cases, the ELM tends to require more hidden neurons than conventional tuning-based algorithm. As a result, the unnecessary or non-optimal hidden biases and input weights may occur.

First, the adjustment of the hidden layer bias and input weights of ELM uses a hybrid technique as so called LCA, and to improve performance of ELM in classifying food items and to find optimal weights of the input layer and the thresholds of the hidden layer an optimisation takes place for the network structure. The original population is made unsystematically, then the algorithm converges toward the optimal weights or the weights that achieve the best accuracy.

A set of hidden biases and input weights composes each member in the population:

$$\theta = [w11,\ w12,\ ...,\ w1K, \ldots\ldots,\ wn2,\ ...\ wnK,\ ...,\ b1,\ b2,\ ...,\ bK] \quad (3)$$

Second, the corresponding output, for each member (a set of weights and biases), are calculated analytically using the MP generalized inverse. Then, an evaluation occurs for the fitness of each member. And the cost function supposedly is root mean squared error.

3. RESULTS AND DISCUSSION

This section discusses the comparative analysis performance of the suggested technique against some existing methods as SVM linear, SVM with RBF and standard ELM based on four measures (including recall, precision, accuracy, and F-measure) using two standard datasets as UEC food 256 [26], and UEC food 100 [27], is conducted as shown in Tables 3 and 4. In addition reports performace of L-ELM depend on accuracy with CNN, DCNN, SVM with MKL and ELM algorithms had been clearly provided in Tables 4 and 5. The programs entirely are run in MATLAB 2015b environment.

In Table 1 the configuration of the algorithm is illustrated. The reason for selecting this configuration is to validate the proposed L-ELM based on the number of neurons equal or higher than the number of inputs. The number of solutions and iterations are selected in the least minimum needed number to obtain good optimisation of ELM.

The results of the comparative analysis of the proposed technique against existing techniques using the UEC food 256 dataset are illustrated in Table 2. The results of the relative analysis of the performance among the methods are presented by using recall, precision, accuracy and F-measure. From the table can be seen that the value of accuracy of SVM linear is smaller than accuracy of ELM, SVM RBF and L-ELM. While, the L-ELM technique is generating the best accuracy in this dataset. It is observed the highest of F- measure (higher recall and precision) is (89%) produced by ELM compared with SVM linear and SVM RBF but is less than value of F-measure for L-ELM technique. Based on the results, the L-ELM performance is more effective than other methods.

Table 3. includes the results of the relative analysis of the performance of the suggested technique against existing techniques using different measures. From the table can be observe that the value of accuracy of SVM linear is smaller than accuracy of ELM, SVM RBF and L-ELM. While L-ELM generates the best accuracy in this dataset. Also, It is noticed the highest of F- measure (higher recall and precision) is (93%) produced by ELM compared with SVM linear and SVM RBF, but is less than value of F-measure is (96%) for L-ELM. The results obtained the L-ELM technique performance is more effective than other methods.

Table 1. Configuration of the proposed algorithm

Dataset	Activation Function	Hidden Neurons	Number of Solutions	Number of Iterations	Number of Classes
UEC Food 256 Dataset	Sigmoid	50	30	10	31
UEC Food 100 Dataset	Sigmoid	50	30	10	30

Table 2. Performance analysis of classification for the proposed technique and other techniques using UEC food 256 dataset

Dataset	Algorithms	Recall	Precision	Accuracy	F-measure
UECFood256 Dataset	ELM	0.889401	0.89636	0.889401	0.892867
	SVM Linear	0.612903	0.529568	0.612903	0.568196
	SVM RBF	0.626728	0.594767	0.626728	0.610329
	L-ELM	0.940092	0.947005	0.940092	0.943536

Table 3. Performance analysis of classification for the proposed technique and other techniques using UEC food 100 dataset

Dataset	Algorithms	Recall	Precision	Accuracy	F-measure
UEC-100 Dataset	ELM	0.933333	0.944497	0.933333	0.938882
	SVM Linear	0.642857	0.563981	0.642857	0.600842
	SVM RBF	0.652381	0.582363	0.652381	0.615387
	L-ELM	0.957143	0.965	0.957143	0.961055

Based on UEC food 256 dataset, the Table 4. demonstrates that the proposed technique yielded better result compared to DCNN [28], with 74% and CNN [29], was produced 80%. Hence, the proposed technique is more appropriate and realistic to use in classifying food items from images. Figure 2. Contains a graphical representation of the outcomes.

Table 4. Results of comparison of the classification for food item images based on accuracy using UEC food 256 dataset

Algorithms	Accuracy
CNN	80.7%
DCNN	74.4%
L-ELM	94.00%

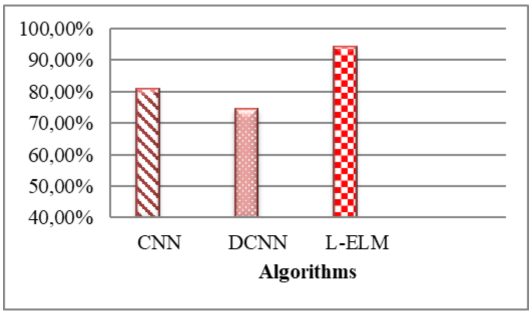

Figure 2. Graphical representation of the classification between L-ELM, DCNN & CNN for UEC food 256 dataset

Table 5. displays the outcomes of the comparative analysis of the proposed technique against several previous studies using UEC food 100 dataset. These studies are [27-31], From the table, it can notice the L-ELM was more realistic in terms of accuracy, which records 95% vs. SVM with MKL [27] with 55%, vs. ELM [31], with 80% and vs. CNN [29], was near of result of L-ELM by 94% also DCNN [30] was 92%. A graphic representation of comparison results is illustrated in Figure 3.

Table 5. Results of the comparison of the classification for food item images based on accuracy using UEC food 100 dataset

Algorithms	Accuracy
SVM with MKL	55.8%
ELM	80.33
DCNN	92.00%
CNN	94.6%
L-ELM	95.71%

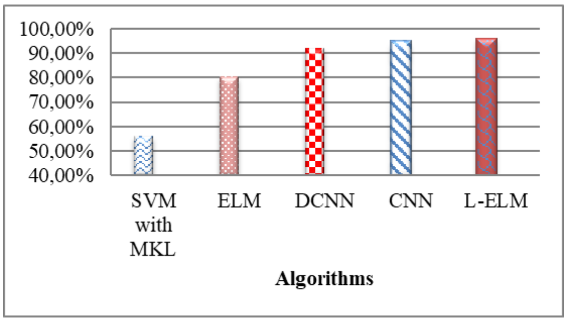

Figure 3. Graphical representation of the classification of L-ELM vs. SVM with MKL vs. DCNN and vs. CNN for UEC food 100 dataset

4. CONCLUSION

This paper presents a current learning algorithm entitled League Extreme Learning Machine which utilizes the merits of LCA and ELM respectively. It makes the use of the rapid minimum norm least-square design to analyze the output weights rather than tuning, and a change form of LCA is used to optimise the hidden biases and the input weights. The suggested method of L-ELM, dissimilar to the gratified-based methods, does not need the activation functions in order to be differentiable, applying L-ELM can be used for training SLENs additional to the nonlinear hidden units like threshold units which are claimed to be easier for hardware implementation. The findings further showed that the experimental output for implementation the L-ELM algorithm was more effective than the other methods.

REFERENCES

[1] Huang, G.-B., Zhu, Q. Y., & Siew, C. K., "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.
[2] Huang, G.-B., & Chen, L., "Convex incremental extreme learning machine," *Neurocomputing*, vol. 70, no.16, pp. 3056–3062, 2007.
[3] Huang, G.-B., & Chen, L. "Enhanced random search based incremental extreme learning machine," *Neurocomputing*, vol. 71, no. 16, pp. 3460–3468, 2008.
[4] Huang, G., Huang, G.-B., Song, S., & You, K., "Trends in extreme learning machines: A review," *Neural Networks*, vol. 61, pp. 32–48, 2015.
[5] Albadra, M. A. A. & S. Tiuna., "Extreme learning machine: a review," *International Journal of Applied Engineering Research*, vol. 12, no. 14, pp. 4610-4623, 2017

[6] Huang, G.-B., et al., "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
[7] Rajesh, R., & Prakash, J. S., "Extreme learning machines-a review and state-of-the-art," *International Journal of Wisdom Based Computing*, vol. 1, no. 1, pp. 35–49, 2011.
[8] Huang, G.-B., "An insight into extreme learning machines: random neurons, random features and kernels," *Cognitive Computation*, vol. 6, no. 3, pp. 376–390, 2014.
[9] Niu, P., Ma, Y., Li, M., Yan, S., & Li, G., "A kind of parameters self-adjusting extreme learning machine," *Neural Processing Letters*, vol. 44, no. 3, pp. 813–830, 2016.
[10] Sun, W., Wang, C., Zhang, C., "Factor analysis and forecasting of CO2 emissions in hebei, using extreme learning machine based on particle swarm optimisation," *Journal of Cleaner Production*, vol. 168, pp. 1095-1101, 2017.
[11] Zhang, X., & Wei, Z., "A hybrid model based on principal component analysis, wavelet transform, and extreme learning machine optimised by bat algorithm for daily solar radiation forecasting," *Sustainability,*vol. 11, no. 15, pp. 4138-4159, 2019.
[12] Eshtay, M., Faris, H., & Obeid, N., "Improving extreme learning machine by competitive swarm optimisation and its application for medical diagnosis problems," *Expert Systems with Application*, vol. 104, pp. 134–152, 2018.
[13] Lu, H., Du, B., Liu, J., Xia, H., & Yeap, W. K., "A kernel extreme learning machine algorithm based on improved particle swam optimisation," *Memetic Computing*, vol. 9, no. 2, pp. 121-128, 2017.
[14] Zhang, Y. et al., "Memetic extreme learning machine," *Pattern Recognition*, vol. 58, pp. 135-148, 2016.
[15] Zhang, X., Yang, Z., Cao, F., Cao, J., Wang, M., & Cai, N., "Conditioning optimisation of extreme learning machine by multitask beetle antennae swarm algorithm," *arXiv preprint*, pp. 1-16, 2018.
[16] Jian, Y., et al., "A novel extreme learning machine classification model for e-nose application based on the multiple kernel approach," *Sensors* (Basel, Switzerland), vol. 17, no. 6, p. 1434, 2017.
[17] Albadr, M. A. A., S. Tiun, F. T. AL-Dhief & M. A. Sammour., "Spoken language identification based on the enhanced self-adjusting extreme learning machine approach," *PloS one*, vol. 13, no. 4, 2018.
[18] Albadr, M. A. A., "Spoken language identification based on optimised genetic algorithm–extreme learning machine approach," *International Journal of Speech Technology*, vol. 22, no. 3, pp. 711-727, 2019.
[19] Albadra, M. A. A. & S. Tiuna., "Extreme learning machine: a review," *International Journal of Applied Engineering Research*, vol. 12, no. 14, pp. 4610-4623, 2017.
[20] Kashan, A. H. "League Championship Algorithm (LCA): An algorithm for global optimisation inspired by sport championships," *Applied Soft Computing*, vol. 16, pp. 171–200, 2014.
[21] Abdulhamid, S. M., Latiff, M. S. A., Madni, S. H. H., & Oluwafemi, O., "A survey of league championship algorithm: prospects and challenges," *Indian Journal of Science and Technology*, vol. 8, no. 3, pp. 101-110, 2015.
[22] Jalili, S., Kashan, A. H., & Hosseinzadeh, Y., "League championship algorithms for optimum design of pin-jointed structures," *Journal of Computing in Civil Engineering*, vol. 31, no. 2, 2016.
[23] Saraswathi, D., & Srinivasan, E., "Mammogram analysis using league championship algorithm optimised ensembled FCRN classifier," *ndones. J. Electr. Eng. Comput. Sci, (IJEECS)*, vol. 5, no. 2, pp. 451-461, 2017.
[24] Xu, W., Wang, R., & Yang, J. An improved league championship algorithm with free search and its application on production scheduling. *Journal of Intelligent Manufacturing*, vol. 29, no. 1, pp. 165–174, 2018.
[25] Kashan, A. H., "League championship algorithm: a new algorithm for numerical function optimisation," *In 2009. IEEE International Conference of Soft Computing and Pattern Recognition (SOCPAR'09)*, pp. 43–48, 2009.
[26] Kawano, Y., & Yanai, K. "Automatic expansion of a food image dataset leveraging existing categories with domain adaptation," *In European Conference on Computer Vision*, pp. 3–17, 2014.
[27] Matsuda, Y., Hoashi, H., & Yanai, K., "Recognition of multiple-food images by detecting candidate regions," *In 2012 IEEE International Conference on Multimedia and Exp*, pp. 25-30, 2012.
[28] Kawano, Y., & Yanai, K., "Foodcam-256: a large-scale real-time mobile food recognitionsystem employing high-dimensional features and compression of classifier weights," *In Proceedings of the 22nd ACM international conference on Multimedia*, pp. 761–762, 2014.
[29] Liu, C., Cao, Y., Luo, Y., Chen, G., Vokkarane, V., & Ma, Y., "Deepfood: Deep learning-based food image recognition for computer-aided dietary assessment," *In International Conference on Smart Homes and Health Telematics*, pp. 37–48, 2016.
[30] Kawano, Y., & Yanai, K., "Food image recognition with deep convolutional features," *In Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 589–593, 2014.
[31] Martinel, N., Piciarelli, C., Micheloni, C., & Luca Foresti, G., "A structured committee for food recognition," *In Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 92–100, 2015.