Fig. 20 demonstrated that the proposed BP/AG had improved the performance of reaching the target error by reducing the number of epochs up to 30%. Again, the proposed CGFR/Ag and BFGS/AG outperformed other algorithms in reaching the target error as can seen in Fig. 21 and 22.

IV. Conclusion

In this paper, a new and improved training method is introduced for fast supervised learning method in the neural network. The performance of the proposed first order and second order methods with adaptive gain (BP-AG, CGFR-AG, BFGS-AG) with standard second order methods without gain (BP, CGFR, BFGS) in terms of speed of convergence evaluated in the number of epochs and CPU time. Based on some simulation results, it's showed that the proposed algorithm had shown improvements in the convergence rate with 40% faster than other standard algorithms without losing their accuracy. It has been shown that the proposed algorithm is also robust as the results have been compared with the 'Matlab neural network toolbox 'implementation. Based on simulation results on selected benchmark datasets, the results clearly show that the proposed method outperforms the standard training algorithms in neural network toolbox. Furthermore, it runs much faster, performs less CPU time, has improved average number of epochs, and better convergence rates without losing their accuracy performance.

Acknowledgment

The authors would like to thank Universiti Tun Hussein Onn Malaysia (UTHM) Ministry of Higher Education (MOHE) Malaysia for financially supporting this Research under Trans-disciplinary Research Grant Scheme (TRGS) vote no. T003. This research also supported by GATES IT Solution Sdn. Bhd under its publication scheme.

References

[1] Thimm G., Moerland F., and Emile Fiesler, The Interchangeability of Learning Rate an Gain in Back propagation Neural Networks. Neural Computation, 1996. 8(2): p. 451-460.

[2] Chin Kim On, Teo Kein Yau, Rayner Alfred, Jason Teo, Patricia Anthony, Wang Cheng, Backpropagation Neural Ensemble for Localizing and Recognizing Non-Standardized Malaysia's Car Plates. International Journal on Advanced Science, Engineering and Information Technology, Vol. 6 (2016) No. 6, pages: 1112-1119.

[3] Eom K. and Jung K., Performance Improvement of Back propagation algorithm by automatic activation function gain tuning using fuzzy logic. Neurocomputing, 2003. 50: p. 439-460.

[4] Holger R. M. and Graeme C. D., The Effect of Internal Parameters and Geometry on the Performance of Back-Propagation Neural Networks. Environmental Modeling and Software, 1998. 13(1): p. 193-209.

[5] Rumelhart D. E., Hinton G. E., and Williams R. J., Learning internal representations by back-propagation errors. Parallel Distributed Processing, 1986. 1 (Rumelhart D.E. et al. Eds.): p. 318-362.

[6] Sotiropoulos D.G., Kostopoulos A.E., and G. T.N., A spectral version of Perry's Conjugate gradient method for neural network training. 4th GRACM Congress on Computational Mechanics (GRACM 2002), 2002. 1: p. 291-298.

[7] Weisss and Kulikowski, Computer systems that learn. 1991: p. 223.

[8] Fahlman S.E., Faster learning variations of back propagation: An empirical study. D. Touretzky, G.E. Hinton and T.J. Sejnowski (editors) Proceedings of the1988 Connectionist Models Summer School, 1988: p. 38-51.

[9] Hollis P. W., Harper J. S., and Paulos J. J., The Effects of Precision Constraints in a Backpropagation Learning Network. Neural Computation, 1990. 2(3): p. 363-373.

[10] Jacobs R.A., Increased rates of convergence through learning rate adaptation. Neural Networks, 1988. 1: p. 295–307.

[11] Kamarthi S. V. and Pitner S., Accelerating Neural Network Training using Weight Extrapolations. Neural Networks, 1999. 12: p. 1285-1299.

[12] Leonard J. and Kramer M. A., Improvement to the backpropagation algorithm for training neural networks. Computer and Chemical Engineering, 1990. 14(3): p. 337-341.

[13] Looney C. G., Stabilization and Speedup of Convergence in Training Feed Forward Neural Networks. Neurocomputing, 1996. 10(1): p. 7-31.

[14] Perantonis S. J. and Karras D. A., An Efficient Constrained Learning Algorithm with Momentum Acceleration. Neural Networks, 1995. 8(2): p. 237-249.

[15] Tollenaere T., SuperSAB: Fast adaptive back propagation wiih good scaling properties. Neural Networks, 1990. 3(5): p. 561-573.

[16] Rigler A. K., Imine J. M., and Vogl T. P., Rescaling of variables in hack propagation learning. Neural Networks, 1990. 3(5): p. 561-573.

[17] Yusuf Hendrawan, Dimas Firmanda Al Riza, Machine Vision Optimization using Nature-Inspired Algorithms to Model Sunagoke Moss Water Status. International Journal on Advanced Science, Engineering and Information Technology, Vol. 6 (2016) No. 1, pages: 45-57.

[18] Shanno D.F., Recent advances in numerical techniques for large scale oprimization. in Neural Nelworks for Control. Miller. Sutton and Werbos, 1990.

[19] Charalambous C., Conjugate gradient algorithm for efficient training of artificial neural networks. IEEE Proc, 1992. 139(3): p. 301-310.

[20] Douglas S. C. and Meng T. H.-Y., Linearized least-squares training of multilayer feedforward neural networks. IEEE lnternational]oint Conference on Neural Networks, 1991. 1: p. 1307-1312.

[21] Wessels L. F. A. and Barnard E., Avoiding false local minima by proper initialization of connections. IEEE Transactions on Neural Networks, 1992. 3: p. 899-905.

[22] Codrington C. and Tenorio M., Adaptive Gain networks. Proceedings of the IEEE International Conference on Neural Networks (ICNN94), 1994. 1: p. 339-344.

[23] Noriega, L., Multilayer Perceptron Tutorial. Lecture notes, 2005: p. 1-12.

[24] Holger R. Maier and Graeme C. Dandy, The effect of internal parameters and geometry on the performance of back-propagation neural networks: an empirical study. Environmental Modelling & Software, 1997. 13: p. 193-209.

[25] Zurada J.M., Introduction to Arti4cial Neural Systems. 1992, St. Paul, MN: West Publishing Company.

[26] Bishop C. M., Neural Networks for Pattern Recognition. 1995: Oxford University Press.

[27] Tai-Hoon Cho, Richard W. Conners, and Philip A. Araman, Fast Back-Propagation Learning Using Steep Activation Functions and Automatic Weight Reinitialization. IEEE International Conference on Systems, Man, and Cybernetics, 1991. 3: p. 1587-1592.

[28] Adrian J. Sheperd, Second Order Methods for Neural Networks-Fast and Reliable Training Methods for Multi-layer Perceptrons, ed. J.G. Taylor. 1997: Springer. 143.

[29] Byatt D., Coope I. D., and PriceC. J., E ect of limited precision on the BFGS quasi-Newton algorithm. ANZIAM J, 2004. 45: p. 283-295.

[30] 28. Fletcher R. and Reeves R. M., Function minimization by conjugate gradients. Comput. J., 1964. 7(2): p. 149-160.

[31] L.Prechelt, Proben1 - A set of Neural Network Bencmark Problems and Benchmarking Rules. Technical Report 21/94, 1994: p. 1-38.

[32] Mangasarian O. L. and W.W. H., Cancer diagnosis via linear programming. SIAM News, 1990. 23(5): p. 1-18.

[33] Erik Hjelmas and P.W. Munro, A comment on parity problem. Technical Report, 1999: p. 1-7.