Applied Mathematics and Computational Intelligence
Volume 8, No.1, Dec 2019 [67-76]

A Human Face Recognition using Alyuda Neurointelligence

Norpah Mahat[1*] and Afifah Sakinah Mohamad Zuki[1]

[1]*Faculty of Computer and Mathematical Sciences, UiTM Perlis, 02600 Arau, Perlis.*

ABSTRACT

Nowadays, face recognition has been one of the most popular studies. It is considered as a highly interesting topic to do a study on. With the advancement of today's technology, face recognition has been used in a wide range of areas. For instance, face recognition is very common in the security industry. The main idea of this study is to identify the best algorithm with the smallest mean squared error (MSE). The analyses were carried out to compare the algorithms with the smallest mean squared error and to improve the previous research on face recognition based on artificial neural networks. The study on face recognition data and their evaluation by neural networks is important in detecting human faces. This study was conducted by using 45 different face images. The architecture for the network was obtained by Alyuda Neurointelligence where the most popular learning algorithm such as Quick Propagation, Conjugate Gradient Descent, Quasi Newton, Limited Memory Quasi Newton, Levenberg-Marquadt, Online Back Propagation and Batch Back Propagation have been implemented and tested to measure the percentage of success. The results indicated that the Adaptive Techniques were extremely useful pattern recognition especially in identifying human faces. The Limited Memory Quasi-Newton becomes the most suitable algorithm to train the human face recognition data with the smallest MSE. Furthermore, this study has shown a strong positive relationship proven by the R-squared and correlation coefficient for all algorithms.

Keywords: Neural Network, Comparative Study, Human Face Recognition.

1. INTRODUCTION

Malaysia is one of the countries that use the security system and authentication. Nowadays the world of information technology makes the security systems becomes very important. The number of systems that have been compromised increasing and verification play a key role as the first line of defense against intruders. The most popular types of verification such as passwords, pin number, card or token and biometric were used widely in the country.

Password is very weak and fragile because of human nature that tends to create passwords that are easily remembered and can be written down elsewhere which make them accessible to others. Cards and tokens can be submitted by anyone and even if the card or token is identified, there is no way of knowing if the person presenting the card is the actual owner. Biometrics is one example that provides a safe method of authentication and identification, as it is difficult to imitate and steal.

Pattern recognition is consisted of handwriting, finger, speaker, voice and face recognition. In recent years, an exploratory research in pattern recognition system using Artificial Neural Network (ANN) methods have been done as it is seen as a method that is successful in terms of pattern recognition [12]. A study of facial recognition is very useful since its potential for commercial applications such as for security systems, film processing, law enforcement, people identification, access control systems and many more. This development of this interest has

*Corresponding Author: *norpah020@uitm.edu.my*

resulted a various technology such as a criminal database system proposed by Nurul Azma et al. [8] where the identification of the suspect is done by face matched.

The face recognition from the images is challenging due to the wide variability of face appearances and the complexity of the image background [7]. There are rapidly increasing needs for information processing and output, due to industrial development, that has new trend and challenges to pattern recognition [9]. However, a reliable performance of computerized facial recognition system still cannot fully achieve. Appearance of the face, head size, orientation and changes in environmental conditions are major changes when the problems arise. One of the fundamental problems in pattern analysis is difficulty making facial recognition.

The purpose of this project is to identify the best methods in ANN in order to recognize a human face. This study aims to identify an algorithm should be used to improve the real face recognition. Other than that, it also wants to identify the software that can give the most stable result.

The main objective of this research is to identify the best algorithm with smallest mean squared error (MSE). The sub objectives are to compare the algorithm with smallest mean squared error (MSE) and to improve the previous research on face recognition based on artificial neural networks.

This paper will focus on testing algorithm for Quick Propagation, Conjugate Gradient Descent, Quasi Newton, Limited Memory Quasi Newton, Levenberg-Marquadt, Online Back Propagation and Batch Back Propagation for human face recognition.

1.1 Artificial Neural Networks (ANN)

Artificial neural networks were evolutionary optimization-based algorithms developed by Ozdemir & Temur [10] and Quin [11]. Neural networks are defined by the neurons and their connections. All neurons are arranged in layers; the sequence of layers defines the order in which the activations are calculated [6].

The neural network was composed of several processors which are very simple and extremely interrelated, called neurons. This is similar to the biological neurons found in the brain. These neurons were connected by links weighted signals from one neuron to another. Neural network learns through repeated adjustments weights. By adjusting the values of connections (weights) between elements a neural network can be trained to perform a specific function [13]. Neural networks commonly were adjusted, or training for a particular input leads to a specific target output. The output need to match the target by adjusting the network based on a ratio of the output and the target [14].

2. RESEARCH METHODOLOGY

Alyuda NeuroIntelligence is a neural networks software application designed to assist neural network, pattern recognition, data mining and predictive modeling experts in solving real-world problems. This software is fast and easy to use and their features only proven neural network modeling algorithms and neural net techniques. It is neural network software for experts, designed for intelligent support in applying neural networks to solve real world forecasting, classification, and function approximation problems [5].

This research focused on training all neural network algorithms in the Alyuda NeuroIntelligence software. The algorithms involved were Quick Propagation, Conjugate Gradient Descent, Quasi Newton, Limited Memory Quasi Newton, Levenberg-Marquadt, Online Back Propagation and

Batch Back Propagation. For its fast processing and wide usage in literature, these algorithms were used for training neural networks on human face recognition data [2],[3],[12],[14].

Based on all of algorithms, this study will compare the algorithms performance with the smallest MSE. The neural network as known has a lot of nodes in all of it layers (at least two or more than two). So the way a neural network works is, when it predicts some value for an output, it compares with the actual output and sends the error back to the nodes. This process is called backpropagation. The mean squared error subsequently applied to calculate the error. There are many other ways to calculate error but the use of mean square error is more valuable in determining the cross-entropy between the distribution of the model predictions and the distribution of the target variable [4].

The Alyuda Neurointelligence software has six main steps to build the ANN architecture [1]. The steps as illustrated in Figure 1 include analyzing data, data pre-processing, design the network architecture, training and testing the networks and validation.

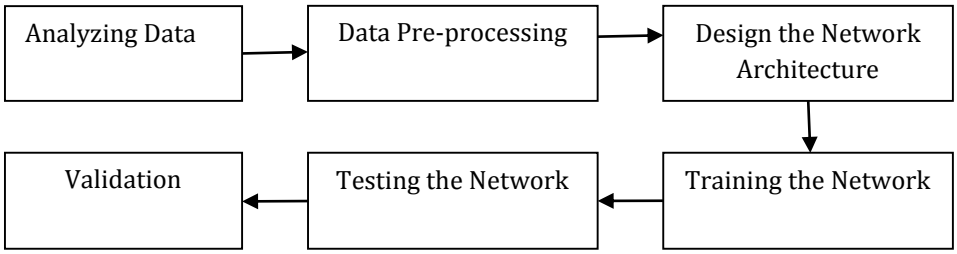

Figure 1. Flowchart using alyuda neurointelligence software.

The first step in training the neural networks was entering the human face data in the Alyuda NeuroIntelligence software. The data from the secondary source consists of 45 different face images [14]. In data analyzing, data anomalies that influence negatively to the network performance should be separated. The anomalies data were divided into two parts known as outliers and missing values. The outliers value that were not within the specified interval while the missing value is an anonymous value that considered as blank cells in the input columns. Data has been imported from file "excel.csv" presents 9 columns and 44 rows were analyzed. Figure 2 presents the data analysis recorded by the system show that 9 columns and 42 rows accepted for neural network training.

It followed by input selection to the respective variables and setting output in the data column. After that, all data will be separated randomly to training, validation and testing set. Hence, the data separation was categorized into 30 data for training set and 6 data for each validation and testing set. The percentage of data partition results was directly generated by the system as 71.43% for training set and 14.29% for validation and testing set (refer to Figure 2).

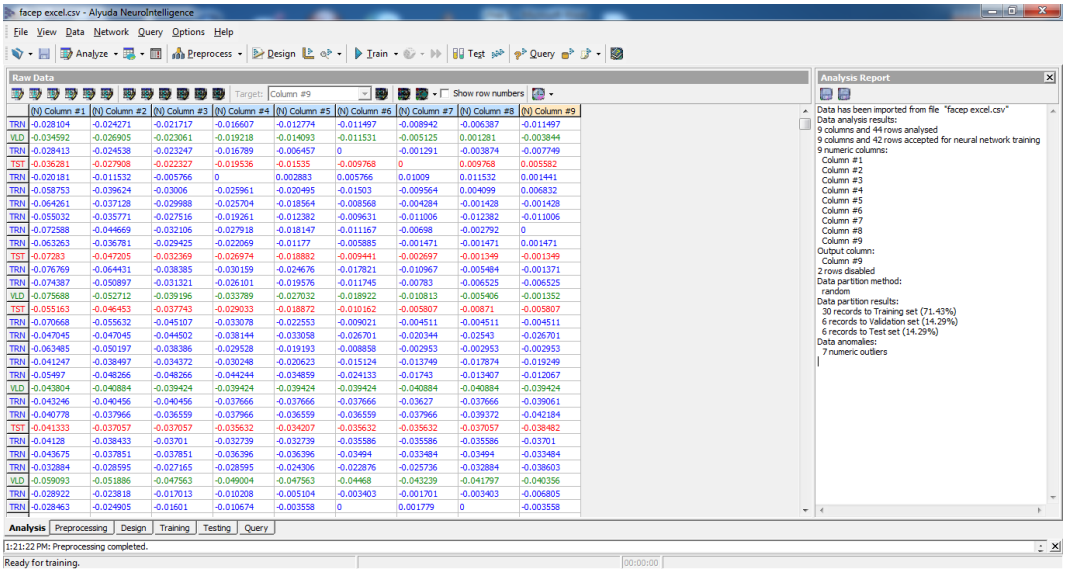

Figure 2. Data analysis.

For data pre-processing, the data in the input columns were normalized to the numerical values. The report shows the columns before pre-processing is 9 and after the pre-processing was staying 9. The input columns scaling indicate the range between -1 to 1 and the output scaling range are between 0 and 1.

For the next step, the networks arhitecture $N^{(8-20-1)}$ was designed by program to ensure the proper variables was taken within the r-squared value 0.94205. The chosen networks architecture consist of one input layer with 8 active neurons, one hidden layer with 20 neurodes and one output layer with one neurode. At this stage the chosen architecture gave the result 0.003673 for train error, 0.005991 for validation error and 0.002648 for test error. The correlation of the best network show a strong positive relationship where the correlation value was 0.975079.

In the training and testing process, the learning and momentum rates were setting at 0.1 and the terminations criteria have been selected to avoid overfitting. The training will stop when the mean square error reach 0.000001 or the process completed at 500 iterations, whichever occurs at the first condition. Also in testing the network, the value of absolute error, network error and mean squared error will be calculated as the best algorithm determination.

Correlation, *r* is a statistical technique used to validate whether and how strongly interrelate of set of variables. The *r* value will be applied in this study as fitting correlation indicator of numerous variables. The network fitting capability is an indication of how well the constructed neural network fits to the topological structure. When calculating the fit between the measured and output values of each data, the r^2 described the fitting effect and the fit line described the fitting curve when $r^2 = 1$[8]. The result of a correlation coefficient (*r*) lies from -1 to +1. When the *r* value closer to 1 show the stronger of the two variables are related. While if the *r* value approaches to 0 indicates no relationship between the dependent and independent variables.

Then, the value of r^2 subsequently determined to analyze the model prediction accuracy that uses mean value for all target. Additionally, if the value is close to 0 means a poor model meanwhile the higher value shows the better the model fits all data.

Figure 3 displays a graph patterns for actual versus ouput result that reflect a good model which is the actual line is located very close to the output line.

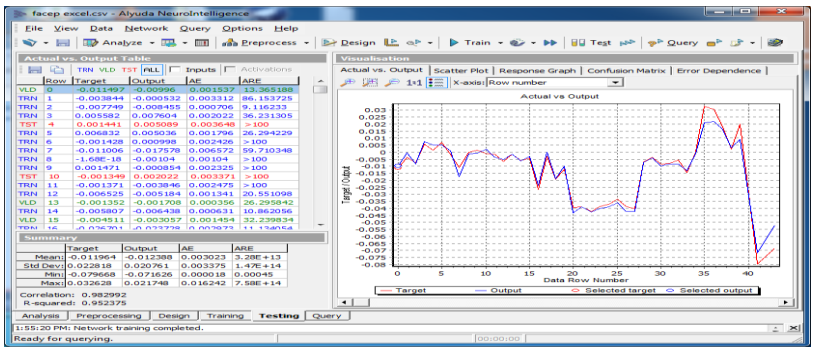

Figure 3. Actual versus output graph.

3. RESULT AND DISCUSSION

In this examination, nine variables selected as predictor variables results from previous inspections. For forecasting performance, the network architecture $N^{(8-1-1)}$ was designed by Alyuda as illustrated in Figure 4. It can be observed that the r^2 value gives 0.94205 which is described a strong relationship between the input data and the corresponding target.

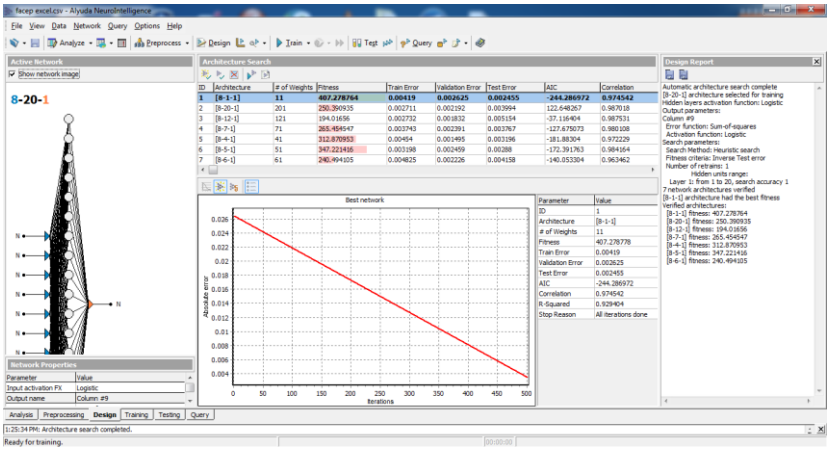

Figure 4. Design the neural network architecture.

Since the neural network architecture has been developed by system, then the learning process was continued with training, testing and training the network. Table 1 presents the error analysis for each algorithm that calculating the absolute errors and network errors.

Table 1 Absolute error and network error for training

Algorithm	Absolute Error	Network Error
Quickprop	0.002200	0.004490
Conjugate Gradient Descent	0.005434	0.001816
Quasi-Newton	0.002349	0.005268
Limited Memory Quasi-Newton	0.002348	0.005060
Lavenberg-Marquadt	0.002447	0.000873
Online Backpropagation	0.003604	0.005963
Batch Backpropagation	0.014508	0.030364

Finally, the powerful of the networks have been examined by using all algorithms in Alyuda. Figure 5 displays the graph of difference value between actual and output for each algorithm after all processes was implemented.

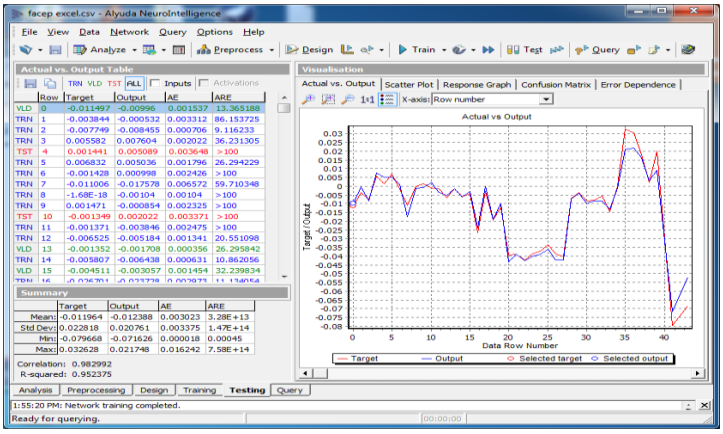

Figure 5. Actual vs output propagation.

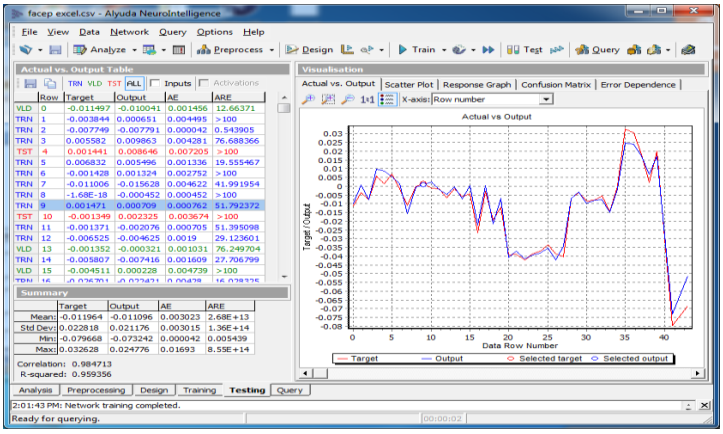

Figure 6. Actual vs ouput for conjugate gradient descent.

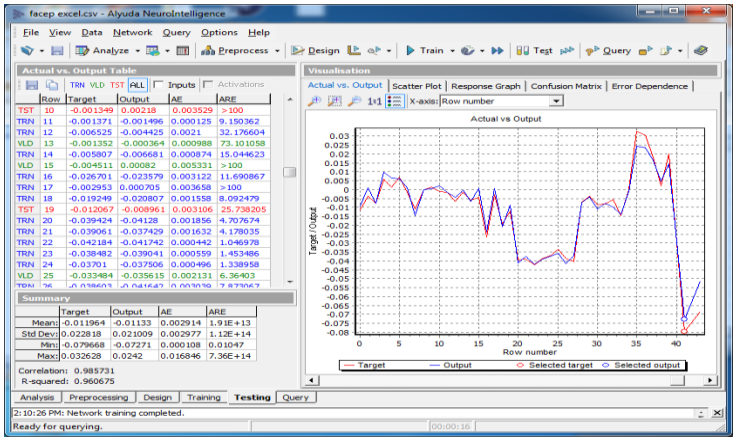

Figure 7. Actual vs ouput for quasi-newton.

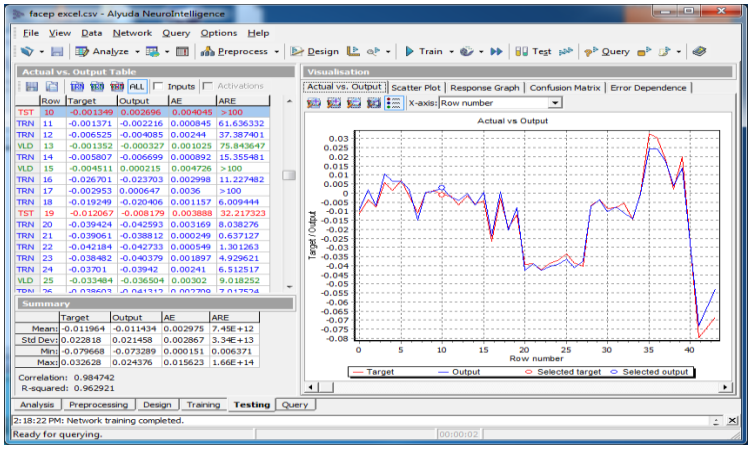

Figure 8. Actual vs Ouput for Limited Memory Quasi-Newton.

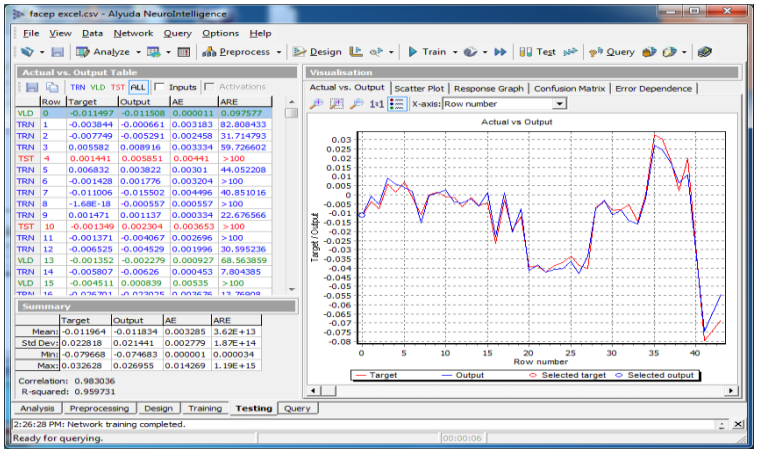

Figure 9. Actual vs Ouput for Levenberg-Marquardt.

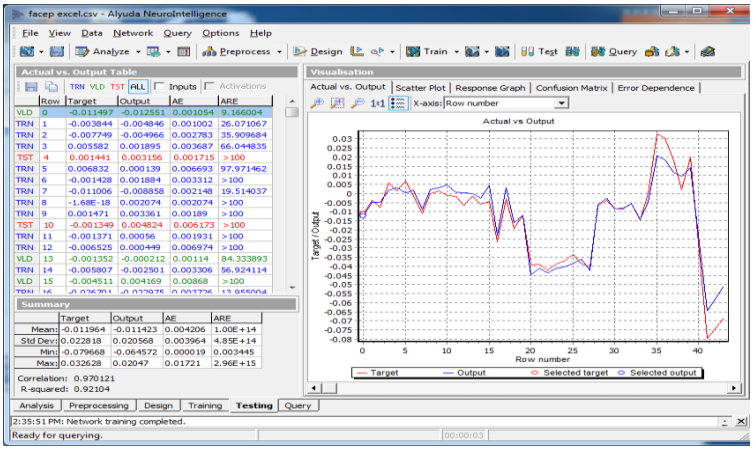

Figure 10. Actual vs Ouput for Online Back Propagation.

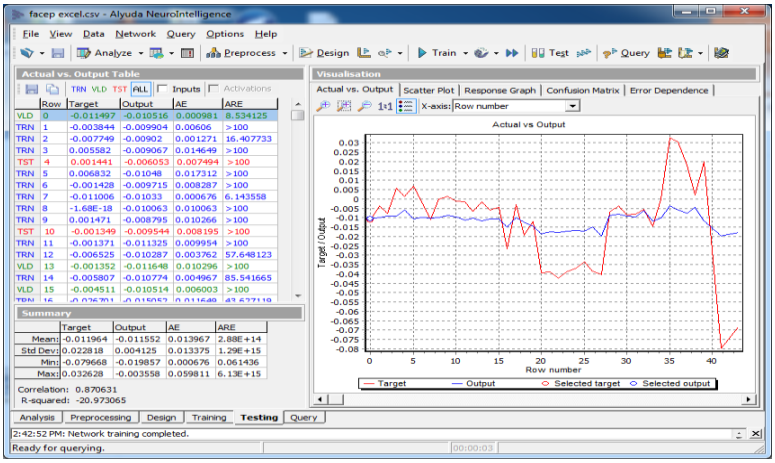

Figure 11. Actual vs ouput for batch back propagation.

After obtaining the results, the algorithms need to find the Mean Square Error (MSE) to determine the best algorithm to conduct a research. By using Microsoft Excel, the MSE has been calculated. The mean squared errors present in Table 2 shows the Limited Memory Quasi-Newton give the lowest MSE value. This means that the Limited Memory Quasi-Newton is the most suitable algorithm to train the human face recognition data.

Table 2 Mean squared error (MSE) for all algorithms

Algorithm	**Mean Square Error (MSE)**
Quickprop	0.0000205226
Conjugate Gradient Descent	0.0000182251
Quasi-Newton	0.0000173571
Limited Memory Quasi-Newton	0.0000170734
Lavenberg-Marquadt	0.0000185126
Online Backpropagation	0.0000334025
Batch Backpropagation	0.0037395600

In terms of successful percentage, over 90% of all algorithms were capable of achieving a good result. Regarding the training phase, the Limited Memory Quasi-Newton algorithm exhibits the best performance of average number of error and percentage of success. The Limited Memory Quasi-Newton also performs a good result.

Furthermore, the Limited Memory Quasi-Newton was more established in the sense of display number of successes and able to achieve a result of over 99%. There were some cases of the solution getting stuck at local minima values for all methods, which means that some of the patterns were not correctly classified.

Table 3 shows the value of correlation and R-squared after the testing process have been done. The results with respect to generalization performance of the tested algorithms indicate that all techniques had proven its capability where the correlation coefficient is close to 1 which means it has a strong positive relationship or positive perfect correlation.

Table 3 Correlation and R-squared results

Algorithm	Absolute Error	Network Error
Quickprop	0.982992	0.952375
Conjugate Gradient Descent	0.984713	0.959356
Quasi-Newton	0.985731	0.960675
Limited Memory Quasi-Newton	0.984742	0.962921
Lavenberg-Marquadt	0.983036	0.959731
Online Backpropagation	0.970121	0.921040
Batch Backpropagation	0.870631	-20.97307

Among all algorithms, the Batch Back Propagation gives a worst performance which is the value of absolute error and network error show greater number among others. A little surprise for that algorithm because the coefficient shows a strong positive relationship but the R-squared show a negative value. The Batch Back Propagation algorithm indicates a larger number of iterations to complete the training. Other algorithms are present R-squared is close to 1 means the variation of output was explained very well by the corresponding target.

4. CONCLUSION AND RECOMMENDATION

In the recent years this research has mostly facial recognition using methods that rely on measurements on the size of the face, rather than neural network techniques. However, these methods were limited in their capabilities but have been proven to work well. Decisions regarding the performance of the tested algorithms indicate that all had proven its capability where the positive correlation coefficient achieved.

In terms of learning speed, the Limited Memory Quasi-Newton and Levenberg-Marquardt algorithms seem to be superior all other training algorithms using fixed parameters in human face recognition. Nevertheless, for the other algorithms, in particular, Quick Propagation, Conjugate Gradient Descent, Limited Memory Quasi Newton, Online Back Propagation and Batch Back Propagation generalize a little better.

For future work, it would also be interesting to investigate further the issue of developing algorithms that will choose an appropriate answer. However, there are lacks of some information about the algorithms in the Alyuda Neurointelligence such as Quasi-Newton algorithm, Limited Memory Quasi Newton algorithm, Online Back Propagation algorithm and Batch Back Propagation algorithm. These algorithms need more study and research on human face recognition to make more people interesting to use the algorithm. This study can encourage people to conduct a research on the other pattern recognition such as voice recognition, lip recognition, handwritten recognition and signature recognition.

REFERENCES

[1] Argyou, A., "Predicting financial distress using neural networks: another episode to the serial". Master thesis, Swedish School of Economics and Business Administration, Helsinki, (2006).
[2] Bryd, R. H., Hansen, S. L., Jorge, N. & Singer, Y., A Stochastic Quasi Newton Method for Large Scale Optimization. SIAM Journal **26**, 2 (2016) 1008-1031.

[3] Cagatay, B., Serdar, D. & Cagdas, H. A. A Comparison of Different Model Selection Criteria for Forcasting Euro/USD Exchange Rates by Feed Forward Neural Network. Int. Journal of Computing, Communications & Instrument Engg. (IJCCIE) **3**, 2 (2016) 271-275.
[4] Ian, G., Yoshua, B., Aaron, C., "Deep Learning-Adaptive Computation and Machine Learning". MIT Press, (2016).
[5] Kim H. Pries & Robert Dunnigan, "Big Data Analytics". CRC Press, Taylor & Francis Group, (2015).
[6] Kumar, J. & Roy, N. A Hybrid Method for Vendor Selection Using Neural Network. International Journal of Computer Applications **11**, 12 (2010) 35-40.
[7] Nandini, M., Bhargavi, P., Raja Sekhar, G., Face Recognition Using Neural Networks. International Journal of Scientific and Research Publication **3**, 3 (2013) 1-5.
[8] Nurul Azma Abdullah, Md. Jamri Saidi, Nurul Hidayah Ab Rahman, Chuah Chai Wen & Isredza Rahmi A. Hamid. "Face Recognition for Criminal Identification: An implementation of principal component analysis for face recognition. The 2nd International Conference on Applied Science and Technology (ICAST'17)", (2017) 1-6.
[9] Oludare Isaac Abiodun, Aman Jantan, Abiodun Esther Omolara, Kemi Victoria Dada, Nachaat AbdElatif Mohamed, Humaira Arshad. State of the art in artificial neural network applications: A survey. Jurnal Heliyon Elsevier **4**, issue 11 (2018).
[10] Ozdemir, D. & Temur, G. T. DEA ANN Approach in Supplier Evaluation System. World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering **54**, 538 (2009) 343-348.
[11] Quin, L. Improving an ANN pruning algorithm based approach to vendor selection. Kybernetes **38**, 3/4 (2009) 314-320.
[12] Wei Li, *et al.,* Using a Backpropagation Artificial Neural Network to Predict Nutrient Removal in Tidal Flow Constructed Wetlands. MDPI open access journal, Water **10**, 1 (2018) 83.
[13] Zainuddin, Z., Mahat, N., Abu Hassan, Y. "Local Adaptive Learning Techniques in Training Artificial Neural Networks", Regional Conference on Integrating Technology in the Mathematical Science, (2003).
[14] Zainuddin, Z., Mahat, N., Abu Hassan, Y., "Improving the Convergence of the Backpropagation Algorithm Using Local Adaptive Techniques", World Academy of Science, Engineering and Technology International Journal of Computer, Electrical, Automation, Control and Information Engineering **1**, 1 (2007) 184-187.