Infrared Face Recognition System Using Cross Entropy Error Function Based Ensemble Backpropagation Neural Networks

Benyamin Kusumoputro and Lina

Abstract—This paper presents the development of cross-entropy error function based ensemble back-propagation neural network and its application as the classifier for frontal face recognition system. As the usually used quadratic error function has a drawback on its local minima problem, the used of a linear cross-entropy error function might directly found the global minima and achieve higher recognition capability. More over, we derive the cross-entropy error function for the negative corelation ensemble backpropagation and compare its characteristics with that of usually used quadratic based ensemble backpropagation neural networks. Experiments on face recognition system are conducted using infrared images, in order to obtain a robust face recognition for security puposes. In these experiments, face images with various expressions were utilized. Results show that the cross-entropy error function based ensemble back-propagation learning system performed better than that of quadratic error function based ensemble back-propagation learning system. It is clearly seen that the cross-entropy error function based ensemble backpropagation system could achieved higher recognition rates and more stable along various data sets, compare with that of quadratic error function based ensemble backpropagation neural network.

Index Terms—Face recognition system, infrared images, ensemble backpropagation neural networks, cross-entropy error function.

I. Introduction

People are able to differentiate and recognize thousands of faces along the entire lifetime. They have the ability to remember and identify hundreds, even thousands, of faces whom they meet in their social lifes. The ability in recognizing those faces still can work well although the faces have changes in certain level; such as age, expressions, illumination and addition of accessories. In this modern era, machines and computers are designed with a smart system so that it can recognize faces as nearly as what human brain does. Today, the technology of face recognition is widely developed and applied, which permits only specific people to access a restricted area.

Automatic face recognition (AFR) is considered to be a natural, a non-intimidating identification method [1], [2]. This

B. Kusumoputro is with the Electrical Engineering Department, Universitas Indonesia, Depok Campus, Depok Indonesia (e-mail: kusumo@ee.ui.ac.id).

Lina is with the Department of Computer Science, Faculty of Information Technology, Tarumanagara University, Jakarta, Indonesia (e-mail: lina@untar.ac.id).

system has been considered to be the leading biometric technology that can also be applied directly into robotic applications. There are many problems, however, mainly due to variations in its lighting condition, which is the most important problem that challenges the robustness of a face recognition algorithms.

Human visual system could ignored the illumination variation when recognizing a person, while the performance of the AFR system decrease significantly when various illumination conditions are present in the input space [3]. To obtain illumination invariant in face recognition, lots of research efforts have been made, and the near infrared (NIR) camera system brings more practical application in real condition. A NIR camera with an additional active NIR lighting provides enough NIR intensity for removing the influence of the environment lighting condition. AFR based on a NIR camera system also less affected by the environment temperature, emotional and health condition of the probe, which can be used both in daytime and nighttime applications.

In this paper, we develop a fast and accurate face recognition system based on active near infrared images for achieving illumination invariant face recognition in indoor on a cooperative-user application. The solution consists of an infrared imaging system, and face recognition algorithms based on ensemble back-propagation neural networks. Artificial neural network with back-propagation algorithm has been studied on researches related to pattern recognition. Training is usually done by iterative updating of weights according to error signal, which is the negative gradient of the quadratic error (QE) function.

The back-propagation learning algorithm using this QE function, however, has a drawback in its possibility of being trapped in local minima. In order to solve this problem, the cross-entropy (CE) error function was proposed as the error signal associated with the output layer is proportional with linear difference between the actual output values with that of the target [4], [5].

The remainder of this paper is organized as follows: Section II describes the back-propagation neural network algorithm using cross-entropy error function, while Section III presents the face recognition system using infrared images. Section IV describe the experimental design and its results, and the conclusion is given in Section V.

II. Cross Entropy Based Ensemble Backpropagation Neural Networks

Neural networks are information processing networks that

mimicking human neural system. Sigmoid activation function which models the actual behaviour of a neuron is mainly used, eventhough there exixts various other functions that represent the characteritics of neuron. Using back-propagation method as a supervised learning, which usually developed based on gradient descent of quadratic error function, has completed the powerful of BP-NN to be a universal approximator. Back-propagation neural networks could approximate any smooth function to an arbitrary degree of accuracy when the tuning parameters could be optimized properly.

Neural network is increasingly used recently to improve the recognition capability of intelligent system because it shows better performance on simulating the non-linear relationship in those problems with the characteristics of parallel computation, self-adapting learning and better fault tolerance [6]-[9]. Neural networks provide a methodology for solving many types of non-linear problems that are difficult to solve by traditional techniques, especially problems without explicitly provide physical or mathematical representation.

The generalization ability of a single neural network, however, depends on many parameters that are difficult to control for different applications in many real world problems. Recently, an ensemble neural network is developed and used as a prominent classifier, since the ensemble neural system utilized learning paradigm where a collection of a finite number of neural networks is trained for the same task [10]-[13].

It originated from Hansen and Solomon's work [14] which shows that the generalization ability of a neural network system can be significantly improved through an ensemble of neural networks, i.e. training many neural networks and then combining their outputs-predictions [13]. Fig. 1 shows the architectural structure of the ensemble BP neural networks. In general, a neural network ensemble is constructed in two steps, i.e. training a number of component neural networks then combining the component predictions by weighted average of the output of each component neural networks [15]-[18].
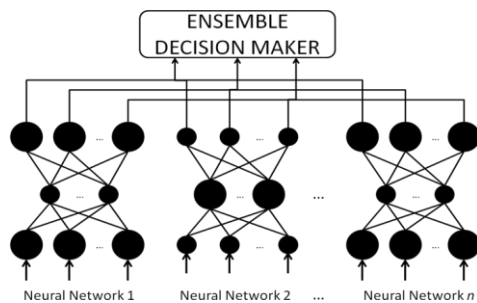

Fig. 1. The ensemble neural of BPNN with negative correlation learning method.

It was found that higher generalization capability of the neural networks is related with the diversity of the component neural ensemble, and further decreasing the generalization error by increasing the diversity of the component neural ensemble [19]. A lot of researches have been tried to propose various learning mechanism of the ensemble neural systems, and yielded numerous methods on how to increase its performances, including Bayesian voting, manipulating the training samples, manipulating the input features, and the

output targets [19].

There are many different types of ensemble learning, i.e., independent training, sequential training and simultaneous training methods [20]. In this paper, however, we will focus on the application of ensemble averaging methods as an independent training method, and the Negative Correlation (NC) learning algorithm as a simultaneous training method. NC learning was firstly introduced in NN ensembles by Yao in [21]-[23] to improve the neural performance in classification, regression, and time-series problems. The idea of NC learning is to introduce a correlation penalty term into the back-propagation error function of each ensemble's component, so that all the networks can be trained simultaneously and interactively [23]. NC learning need to perform periodic communication between component networks for determining the optimized propagated error for changing the weights of each component.

Supervised learning of a neural networks can be viewed as a curve-fitting process. A training vector pairs that consists of an input vector from an input space and a target vector as the neural response are presented to the network. Using the defined learning algorithm, the neural networks performs the adjustments of its weights so that the error between the actual output vectors and the target vectors is minimized relative to some optimization criterion. Once trained, the neural network performs the interpolation in the output vector space which then referred as the generalization capability. Consider a BP NN with the input layer consists of neurons x_n, $n = 1, ..., N$, a hidden layer that consists of neurons z_m, $m = 1, ..., M$, and an output layer that consists of neurons y_l, $l = 1, ..., L$.

Let $X(n) = \{x(1), t(1), x(2), t(2), ..., X(N), t(n)\}$ represents the training data where N is the number of patterns and $\{x, t\}$ are the input vector $x \in R^n$ and target-output vector $t \in R^m$, respectively. Suppose we form an ensemble of I networks, such as in Fig. 1, and the ensemble joint output Y is a simple averaging of the i-th network outputs $Y_i, \forall_i = [1, 2, ..., I]$. The ensemble joint error E can also be written as a simple averaging of the i-th network error. However, instead of the usually used a quadratic error function [10, 24], $E_i(n) = \|t(n) - Y_i(n)\|^2$; we would like to derive the ensemble error by using Cross Entropy (CE) error function such as

$$E_i(n) = -t(n)\ln(y_{im}(n) - (1 - t(n))\ln(1 - y_{im}(n) \qquad (1)$$

thus Y and E are given by

$$Y(n) = \frac{1}{I}\sum_{i=1}^{I} y_{im}(n); \; E(n) = \frac{1}{I}\sum_{i=1}^{I} E_i(n) \qquad (2)$$

where $n = 1, 2, ..., N$ denote the index of training sample.

Define a correlation penalty P_i that measures the error correlation between the i-th networks and the rest of the networks in the ensemble [21], [22] are as follows,

$$P_i(n) = (y_{im}(n) - Y(n)\sum_{\forall j \neq i} y_{jm}(n) - Y(n) \qquad (3)$$

with $y_{im}(n) - Y(n)$ is the error of the i-th network, and $\sum_{\forall j \neq i}(y_{jm}(n) - Y(n))$ the joint error of the rest of the networks in

the ensemble. By combining the rest of the component networks into the *i*-th network error calculation, the new error function E_i is then a weighted sum of the original error function and the penalty function P_i, given by

$$E_{im}(n) = -t_{im}(n)\ln(Y_{im}(n)) - (1 - t_{im}(n))\ln(1 - Y_{im}(n)) + \lambda P_i(n) \quad (4)$$

where $0 \le \lambda \le 1$ is the hyper parameter that adjusts the strength of the correlation penalty term.

The process of learning for updating the weights of the *i*-th network is then calculated through partial derivative of the ensemble error E with respect to Y_i [11], [15], using Eq. (1)-(3):

$$\frac{\partial E_i(n)}{\partial y_{im}(n)} = \frac{-t_{im}(n)}{y_{im}(n)} + \frac{(1 - t_{im}(n))}{(1 - y_{im}(n))} + \lambda \frac{\partial P_i(n)}{\partial y_{im}(n)}$$

$$= \frac{y_{im}(n) - t_{im}(n)}{y_{im}(n)(1 - y_{im}(n))} + \lambda \sum_{j \ne 1}^{s} (y_{jm}(n) - Y(n))$$

$$= \frac{y_{im}(n) - t_{im}(n)}{y_{im}(n)(1 - y_{im}(n))} + \lambda(IY(n) - y_{im}(n) - (I-1)Y(n))$$

$$= \frac{y_{im}(n) - t_{im}(n)}{y_{im}(n)(1 - y_{im}(n))} + \lambda(Y(n) - y_{im}(n)) \quad (5)$$

As for comparison, the weights updating of the *i*-th network calculated through the usually used Quadratic Error (QE) function is

$$\frac{\partial E_i(n)}{\partial y_{im}(n)} = y_{im}(n) - t_{im}(n) + \lambda \frac{\partial P_i(n)}{\partial y_{im}(n)}$$

$$= y_{im}(n) - t_{im}(n) + \lambda \sum_{j \ne 1} y_{jm}(n) - Y(n)$$

$$= y_{im}(n) - t_{im}(n) + \lambda(IY(n) - y_{im}(n) - (I-1)Y(n))$$

$$= (1 - \lambda)(t_{im} - y_{im}) + \lambda(t_{im} - Y(n)) \quad (6)$$

III. FACE RECOGNITION UNDER VARIOUS ILLUMINATION CONDITIONS

A face recognition system usually consists of five stages, i.e. data collection, pre-processing, feature extraction, training and testing. Most current face recognition systems are based on face images captured by visible light camera system [1], [2], [7], [8]. To achieve reliable results, face recognition should be performed based on intrinsic factors of the face, mainly to the reflectance of the facial surface. Extrinsic factors such as hairstyle, posture, expression, including glasses which make the distributions of the face data more highly complex should be minimized. Among the extrinsic factors, problems with the uncontrolled environmental lighting are the most difficult issues to solve for a face recognition system in practice.

We reported here a face recognition system using a near infrared camera system as a sensory system and a cross-entropy based ensemble back-propagation neural network as a classifier. We build an active near infrared imaging system which is able to produce front-lighted face

images regardless of visible lights in the environment.

In order to obtain a database that can be used for some validation of our face recognition algorithms, we performed a series of face acquisition sessions in our department, and named it as IR-UI (infrared-UI) database. The near infrared image database used in the experiments are captured using a LYD 801C Infrared Video camera with 480×640 pixels. Data acquisition was set in office environment with completely dark condition. This database consists of 20 persons, however only 10 persons are utilized at present; which is depicted in Fig. 2.

The frontal face acquisitions were conducted as follows: the subject sits in a special chair within a distance of 1 m and just before the camera was put in the recorded scheme, the environment lightning is turned off. A frontal face database with slightly variations within 5° were allowed and variations of pose and facial expressions between the different records were exists in the gallery.

In the experiments, 200 face images with 20 images for each person in the database gallery were used. All face images were cropped into 40 pixels height and 30 pixels wide according to their eye coordinates, and the database was divided into two data sets, the training and the testing data sets.

Fig. 1 shows the example of various faces and its expressions captured by an infrared camera system. Since these images are represented in a very high dimensional space, they could not be directly applied to face recognition algorithms due to efficiency reasons. PCA was then used to efficiently represent a collection of images by reducing their dimensionality.

Fig. 2. Sample of infra red face image with various expressions.

Feature Extraction Using Principal Component Analysis

The main concept of Principal Component Analysis (PCA) is to reduce the dimensions of the images in the database to a new space with a smaller number of dimensions, yet without loss of significant information. It can be achieved by transforming the data set to a new variable set by Principal Components (PCs). PCs comprise some chosen variables which are already put in order [11]. PCA maps the original image in *N*-dimensional space onto an *h*-dimensional feature space (called the Eigen space), where normally $h \ll N$.

In the training stage, first the images are normalized in the form a vector $x(p)$, $p = 1, ..., P$ the number of overall face images, with N the number of pixels of each image, that can be written as:

$$x^N(p) = [x_1, x_2, ..., X_N]^T \quad (7)$$

By subtracting the average image C of all images, the learning set W can be obtained using:

$$W = [x_1 - c, x_2 - c, ..., x_N - c]^T \quad (8)$$

Next, the covariance matrix is defined by:

$$Q = WW^T \quad (9)$$

and the eigenvector e_i and the corresponding eigenvalues λ_i are determined by solving the eigenvector decomposition problem:

$$\lambda_i e_i = Qe_i \quad (10)$$

A threshold value S is then used for dimension-reduction:

$$\frac{\sum_i^h \lambda_i}{\sum_i^N \lambda_i} \geq S \quad (11)$$

where $h << N$.

Finally, the projection of image $x^N(p)$ from image space with N dimension as $x^h(p)$ in the Eigen space with h dimension is obtained by:

$$x^h(p) = [e_1, e_2, ..., e_h]^T x_n - c) \quad (12)$$

IV. EXPERIMENTS AND RESULTS

It is well known that back-propagation algorithm comprises with numerous parameter which are usually difficult to determine. In the preliminary experiment, some parameters are tuned in order to have the best performance of the neural networks. Ten times experiments are conducted conceutively, and after carefully studied, the parameters are determined as follows. The eigenspace are defined by 50 dimensions.

We have decided to used two data sets to analyze the performance characteristics of the neural networks according to the image dimensions. The first data set, Dataset#1, is composed of 40×30 pixels while the second data set, Dataset#2, is composed of 30×30 pixels. Those data set are derived directly from the infrared camera

TABLE I: RECOGNITION RATE COMPARISON OF THE CE BASED BP-NN AND THE QE BASED BP-NN FOR TWO TYPE OF DATASETS

Dataset	Training Data		Testing Data	
	QE Function	CE Function	QE Function	CE Function
Dataset#1	100	100	94.7	98.6
Dataset#2	100	100	91.1	97.5

TABLE II: RECOGNITION RATE COMPARISON OF THE CE BASED AND THE QE BASED ENSEMBLE BPNN

Neural Component	Dataset#1		Dataset#2	
	QE Function	CE Function	QE Function	CE Function
3 Components	96.6	99.4	94.4	98.0
5 Components	96.8	99.8	94.7	98.1

PCA is then used to reduce the dimension of the face image in each of the data sets, i.e., 1200 dimensions for images in Dataset#1 and 900 dimensions for images in Dataset#2, which is to be projected onto an eigenspace with only 50 dimensions. The hidden neurons is defined by 30, in order to reduce significantly the computational cost that related directly to the number of the hidden neurons.

Learning rate (α) is determined as 0.4 for speeding up the convergence speed, while the momentum learning rate (μ) is determined to be 0.8 for the same reason. For each of the determined data, the experiments are conducted ten times conceutively, and the average values of the recognition rate are recorded. Table I shows the experimental results using training-testing paradigm of 50:50, for each of Dataset#1 and Dataset#2, respectively.

As can be seen from this table, the recognition rate for the training phase for both datasets using QE based BP-NN and CE based BP-NN show the highest rate when its convergence. Please note that the computational cost for both QE and CE methods are comparable, with CE based BPNN shows just a little bit lower. However, it is clearly seen from this table that in the testing phase, the CE based BP-NN have showed higher recognition capability.

For both QE based BP-NN and CE based BP-NN, the recognition rate for the Dataset#2 are lower compare with that of Dataset#1. Please note that eventhough the number of the used principal component is the same for both datasets, the real dimensions of the face images in the Dataset#1 are higher that that in the Dataset#2. Lower recognition rate of the BP-NN when it is used in Dataset2 is then according to this difficulties.

When the recognition rate comparison of the QE based BP-NN and the CE based BP-NN for the Dataset#1 show a comparable results, however, the difference of its recognition rate have become higher for the Dataset#2. From this experiments, we have seen that the CE based BP-NN shows more stable recognition capability, i.e., a comparable recognition rate when it is used to recognize the Dataset#1 and the Dataset#2. However, the QE based BPNN do not have a stable recognition capability such as in the QE based BPNN.

The recognition rate of the QE based BPNN show a significant lower recognition capability when it is used for Dataset#2. It is clearly seen from this comparison that, the QE based BP-NN has a higher margin of error compare with that of the CE based BP-NN, showing that the CE based BP-NN has more stable generalization characteristics.

Comparison of CE and QE Based Ensemble BP-NN

As explained earlier in this paper, the ensemble neural networks consist of several single neural network component trained simultaneously by back-propagation algorithms. Table II shows the recognition rates of the ensemble neural network with negative correlation learning method, for the two datasets that was the same as in the experiments with single BPNN. As can be seen in this table, the recognition rate of the ensemble neural networks using 50% of the learning-dataset for cross-entropy error function are always higher compare with that of quadratic error function.

Experiments results also shown that recognition rate for Dataset#2 are always lower compare with that of Dataset#1,

show that the Dataset#2 is more difficult problem. This finding is in accordance with the results on single BPNN as can be seen in Table I. As also depicted in both tables, the recognition rate of the ensemble BPNN are always higher compare with that of single BPNN. The more component utilized in ensemble system, the higher recognition rate of the system.

As clearly seen in Table II, the recognition rate of the ensemble BPNN with 5 neural elements are always higher compare with that of 3 neural elements. This results show that increasing the diversity of the component neural ensemble increase the recognition capability of the system. However, this slight increment of recognition capability also increases the computational cost of the neural system. It is recommended that using 3 neural components is the best trade-off betwen the increment recognition capability and its computational cost.

V. CONCLUSION

In this paper, we presented our developed face recognition system using an infra-red camera that could be used in recognizing face under video surveilance. Near infra-red camera with additional active infra-red lighting provides enough NIR intensity for removing the influence of the environment lighting condition, hence providing a face recognition system that is very robust to the variance of the environment lighting, even in a very dark light condition. We have carried out extensive experiments to evaluate the performance of the ensemble Back-propagation neural networks using two types of difference error functions. Instead of the usually used quadratic error function, a cross-entropy error function is utilized. A comparison performance analysis are conducted, and results show that the CE based BPNN performed higher recognition rate and more stable compare with that of QE based BPNN, for all of the used datasets. The superiority of the CE based BPNN is also justified in the ensemble form of BPNN. The ensemble neural networks gives higher recognition rate compared with that of the single neural networks, with the highest recognition rate is achieved when using 5 neural components. However, considering this ensemble configuration has a drawback on its higher computational costs, the optimal ensemble configuration in our case is an ensemble with 3 neural components. Since the number of persons in the database that was used in those experiments are very limited, detailed comparison with other type of neural networks such as radial basis function and Lavenberq-Marquard based multilayer perceptron are still impractical at present, and are left for further work.

REFERENCES

[1] J. Ashbourn, *Biometric: Advanced Identity Verification*, Berlin: Springer Verlag, Heidelberg, New York, 2002.
[2] G. J. Ortega, J. Bigun, D. Reynolds, and R. J. Gonzales, "Authentification gets personal with biometrics," *IEEE Signal Processing Magazine*, vol. 21, no. 2, pp. 50-62, 2004.
[3] W. Zhao, R. Chellapa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399-458, 2003.
[4] V. A. Ooyen and B. Nienhuis, "Improving convergence of the back-propagation algorithm," *Neural Networks*, vol. 5, pp. 465-471, 1992.
[5] M. D. Richard and R. P. Lipmann, "Neural networks classifier estimate Bayesian a posteori probabilities," *Neural Computation*, vol. 3, pp. 461-83, 1991.
[6] V. Jothiprakash, R. B. Magar, and S. Kalkuthil, "Rainfalls-runoff models using adaptive neuro-fuzzy inference system (ANFIS) for an intermittent river," *Int. J. Artif. Intell.*, vol. 3, no. A09, pp. 1-23, 2009.
[7] A. Garcia, A. Luviano-Juarez, A. Chaires, and T. Poznyak, "Projectional dynamic neural networks identifier for chaotic systems: applications to Chua's circuit," *Int. J. Artif. Intell.*, vol. 6, no. S11, pp. 1-18, 2011.
[8] E. Joelianto, S. Widyantoro, and M. Ichsan, "Time series estimation on earthquake events using ANFIS with mapping function," *Int. J. Artif. Intell.*, vol. 3, no. A09, pp. 37-63, 2009.
[9] K. K. Sarma, "Neural network based feature extraction for assamese character and numeral recognition," *Int. J. Artif. Intell.*, vol. 2, no. S09, 37-56, 2009.
[10] J. M. Zurada, *Introduction to Artificial Neural Systems*, West Publishing Company, Saint Paul, Minnesota, 1992.
[11] J. G. Carney and P. Cunningham, "Confidence and prediction intervals for neural network ensembles," in *Proc. the Int. Joint Conf. on Neural Networks*, Washington DC, 1999, vol. 2, pp. 1215-1218.
[12] J. G. Carney and P. Cunningham, "Tuning diversity in bagged neural networks ensemble," Technical Report TCD-CS-1999-44, Trinity College, Dublin, 1999.
[13] P. Sollich and A. Krogh, "Learning with ensembles: how over-fitting can be useful," in *Advances in Neural Information Processing Systems*, D. S. Touretzky, M. C. Mozer, and M. E. Hasselmo, Eds. MIT Press Cambridge, Massachussetts, 1996, pp. 190-196.
[14] L. K. Hansen and P. Salomon, "Neural network ensembles," *IEEE Trans. PAMI*, vol. 12, no. 10, pp. 993-1001, 1990.
[15] S. B. Cho, J. H. Ahn, and S. I. Lee, "Exploiting diversity of neural ensembles with speculated evolution," in *Proc. the Int. Joint Conf. on Neural Networks*, Washington DC, vol. 2, 2001, pp. 808-813.
[16] Y. Jiang, Z. H. Zhou, and Z. Q. Chen, "Rule learning based in neural networks ensemble," in *Proc. the Int. Joint Conf. on Neural Networks*, Honolulu, Hawaii, vol. 2, 2002, pp. 1416-1420.
[17] M. P. Perrone and L. N. Cooper, "When networks disagree: Ensemble methods for hybrid neural networks," in *Neural Networks for Speech and Image Processing*, R. J. Mammone, Ed. London, 1993.
[18] D. Jimenez and N. Walsh, "Dynamically weighted ensemble neural networks for classification," in *Proc. the Int. J. Conf. on Neural Networks Anchorage*, Alaska, 1998, vol. 1, pp. 753-756.
[19] A. Krogh and J. Vedelsby, "Neural networks ensembles, cross validation and active learning," *Advances in Neural Information Processing Systems*, vol. 7, Cambridge, MA, 1995.
[20] M. M. Islam, X. Yao, and K. Murase, "A constructive algorithms for training cooperative neural networks ensembles," *IEEE Trans. Neural Networks*, vol. 14, no. 4, pp. 820-834, 2003.
[21] Y. Liu, *Weights Updated Voting for Ensemble of Neural Networks Based Incremental Learning*, Springer-Verlag, Part I, LNCS 5551, 2009, pp. 661-669.
[22] Y. Liu and X. Yao, "Negatively correlated neural networks can produce best ensembles," *Australian J. of Intell. Inform. Process. Syst.*, vol. 4, no. 3, pp. 176-185, 1997.
[23] Y. Liu and X. Yao, "Ensemble learning via negative correlation," *Neural Networks*, vol. 12, pp. 1299-1404, 1999.
[24] L. Breimen, "Combining predictors," in *Combining Artificial Neural Nets: Ensemble and Modular Multinet Systems*, A. J. C. Sharkey, Ed. New York, 1999, pp. 31-50.
[25] J. Lu and Y. P. Tan, "Uncorrelated discriminant nearest feature line analysis for face recognition," *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 185-189, 2010.
[26] X. Chen, P. J. Flynn, and K. W. Bowyer, "Infrared and visible light face recognition," *J. of Comp. Vision and Image Understanding*, vol. 9, pp. 332-358, 2005.
[27] B. Kusumoputro, Lina, and B. Kresnaraman, "Face recognition system of infrared images using ensemble backpropagation neural networks," *Int. J. Arti. Intell.*, vol. 7, no. A11, pp. 401-416, 2011.
[28] I. J. Joliffe, *Principal Component Analysis*, Second Edition, Springer, 2002.

Benyamin Kusumoputro was born in Bandung, West Java, Indonesia on November 17, 1957. He finished his undergraduate course from the Department of Physics, Bandung Institute of Technology, Bandung, Indonesia in 1981. He entered the Graduate School of Optoelectronics in University of Indonesia and graduated in 1984. In 1990 he passed the doctoral course at Department of

Electrical and Electronic Engineering, Graduate School of Science and Engineering, Tokyo Institute of Technology. He was granted a doctor of engineering degree in 1993. He is now a professor of computational intelligence at the Department of Electrical Engineering, University of Indonesia. His research of interest include the 3D modelless face recognition system, intelligent control of unmanned aerial vehicle and quality classification of nuclear fuel using computational intelligence methods. He received Best Researcher Award from Indonesia Ministry of Reseach and Technology in 2006.

Lina was born in Padang, West Sumatra, Indonesia on 20 February 1979. She finished her under graduate course from Department of Computer Science, Faculty of Engineering, Tarumanagara University, Jakarta, Indonesia in 2001. She then entered the Graduate School of Computer Science at University of Indonesia and graduated in 2004. In 2005 she was a research student at the Department of Media Science, Graduate School of Information Science, Nagoya University, and entered the doctor course in 2006 and granted a Ph.D degree in 2009. She is now a lecturer at the Department of Computer Science, Faculty of Information Technology, Tarumanagara University. Her research interests are in the area of image recognition and intelligent system.