

Thai Buddhist Amulet Recognition System

Waranat Kitiyanan, and Chomtip Pornpanomchai

Abstract—The objective of this research is to recognize digital images of Thai Buddhist amulets by using image processing techniques and artificial neural network (ANN). The scope is to recognize 100 kinds of powder Buddhist amulet. To extract features from collected samples, the extraction process considers on shape, color, and texture. In total, 54 feature values are calculated fed into recognition process. To recognize those features, there are two ANN models applied, which are perceptron and multi-layer perceptron. In the experiment, those two models are tested with a test set. On average, the perceptron can recognize the amulet images correctly at 97.67 percent, and the multi-layer perceptron can produce the best classification result at 100 percent. The average computational time for image preprocessing and feature extraction step is about 7.18 seconds per image. To proceed recognition step, the system approximately takes 0.19 second per image for the perceptron and 0.18 second per image for multi-layer perceptron on average.

Index Terms—Buddhist amulet, pattern recognition, image processing, artificial neural network.

I. INTRODUCTION

In Thailand, there are more than 90 percent of population who are Buddhists. Thailand and Buddhism are firmly related in many aspects. It has been influencing the Thai history, art, and culture for a long time. There are at least 25,000 Buddhist temples established but, they cannot connect Thai people and Buddhism together without other two things, which are monks and ceremony. Mostly, Thai Buddhist people believe that attending to ceremony led by monks comes up with luck and good things. After joining a ceremony, they may get a small Buddha image believed that it will bless and protect them from bad things as amulet.

Buddhist amulets are made in many types and styles. They have different size, shape, material, and the appearance of Buddha. It makes them unique and attractive to collectors who have faith and love to see Buddhist art. Because of the demand, Buddhist amulet becomes a trading community. As a result, amulet trader is a well-known occupation in Thailand.

Nowadays the trading is not only on the street, but it is on the internet also. Traders share their collections by using digital media rather than physical media such as amulet magazine. The traders like to post their collection on the internet via online services such as, internet forum, and social network, so they may get some interesting offers as well. In addition, search engines can find several results of Buddhist amulet images if there are some related keywords used. It indicates that putting amulets into digital form for trading is a

usual way for traders to get trading easier and faster.

Basically, the traders have many kinds of Buddhist amulet, so the skill they need is to remember them. It is very important because traders have to manage their collection or share necessary information requested by other traders. Lacking of ability to remember those amulets is a problem for traders because they need to check the offered amulets if they the same as their own collections. Moreover, for new collectors and traders, it is hard to check that the amulet's information is corresponding to the posted amulet's image. It is almost impossible to know all of amulet kinds. Technically, traders will have a trading only when they can verify the amulet kind and genuineness.

In the area of artificial intelligence, this kind of problem can be solved by using pattern recognition technique. A popular approach is called artificial neural network (ANN). ANN is a supervised learning machine inspired by the biological nervous system. It is used to classify data of digital images in several recognition experiments.

By the reasons, this research is motivated to invent a system, which is able to mitigate the problems of Buddhist amulet management inefficiency and non-experience trader. The goal is to collect a sample dataset of Buddhist amulet images then design, implement, and evaluate a recognition system using techniques of digital image processing and artificial neural network for classifying those amulet images. Additionally, the system has to be designed for easy use and acceptable result as well.

II. LITERATURE REVIEWS

Actually, there are some researchers tried to develop Buddhist amulet systems. Even they had different motivations, their systems perform a common functionality, which is to identify types or styles of Buddhist amulet. The other correspondence is that they all involved with digital image processing and artificial intelligence. However, their techniques differed from each other. This section will discuss about them respectively.

A. *A Robust Coin Recognition Method with Rotation Invariance*

P. Thumwarin et al. (2006) had an intention to build a system for recognizing Thai currency coins. The point is the dataset they used. It did not contain only the images of currency coins, but it had amulet-coin images also. The technique they used to extract features was to collect grey-level intensities from many circles in different radii of a coin and analyze them with Fourier transform. The coefficient values are rotation-invariant features that can be matched by using Euclidean distance equation. For the test, they compared 69 coins to a reference amulet coin, and the system could classify between matched and mismatched type for 100 percent [1].

Manuscript received November 9, 2014; accepted November 12, 2014

Waranat Kitiyanan is a master student of the faculty of Information and Communication Technology, Mahidol University, Bangkok, 10120 Thailand (e-mail: wrnt.058@gmail.com).

Chomtip Pornpanomchai is an Assistant Professor in of the faculty of Information and Communication Technology, Mahidol University, Bangkok, 10120 Thailand (e-mail: chomtip.por@mahidol.ac.th).

B. Identification of Amulets with Special Feature Matching

C. Korbuakaew (2007) presented a recognition system for powder Buddhist amulet. His motivation came from a willing to conserve Thai Buddhist art. For the scope, his dataset contained only images of powder amulet of Buddha. To extract special features from amulet images, he used the technique of canny edge detection to proceed. 140 images were used as archetypes that input images can refer to. For matching process, he used correlation coefficient to perform. According to his experiment, the system's accuracy rate is 93.5 percent [2].

C. Buddhist Amulet Recognition System

J. Wongkorsub et al. (2010) introduced Buddhist amulet recognition system (BARS). The system is dedicated to recognize amulet coins. 53 types of coins are collected in there dataset. To proceed feature extraction, the gray-scale conversion was used. The gray-scale values of each pixel were treated as features. The difference of gray-scale intensities between an input image and reference images were calculated to find correlation of them. Low difference value represented high correlation. 80 percent of testing data were classified correctly [3].

D. Buddhist Amulet Coin Recognition by Genetic Algorithm

N. Srisupornwattana et al. (2013) published Buddhist amulet coin recognition by using genetic algorithm. Mainly, the researchers considered two things as features, which are texture and edge. They applied canny edge detection to trace amulet's edges and grey level co-occurrence matrix (GLCM) to analyze amulet's texture. In matching phase, the system used a combination rule-based technique and genetic algorithm to classify their input images. 130 styles of amulets were tested, and 91.53 percent of them are matched [4].

According to the referred researches, the researchers implemented systems that were able to recognize Buddhist amulet with various techniques. Their result were also acceptable. However, there was no one who tried to apply ANN with this problem.

III. METHODOLOGY

A. Hardware and Software Specification

The pictures in this work are taken by using mobile phone's camera (Sony Xperia M Dual). For the computer, it will perform based on following specifications: 1) CPU Intel® Core™ i5-2410M Processor (3M Cache, up to 2.90 GHz), 2) Memory DDR3 4GB, and 3) Hard disk 640 GB. The needed software in implementation is MATLAB R2013a (The Math Works Inc.; Natick, MA, USA) and Microsoft Windows 7 (Microsoft Corp.; Redmond, WA, USA) as the operating system.

B. System Design

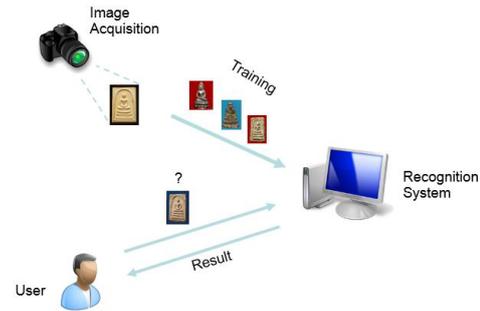


Fig. 1 System Overview.

Briefly, the system overview is described in the Fig. 1. The images of amulet will be taken by image acquisition process. Then, those images is put into the recognition system for training. After training is done, the user can input an image of unknown amulet to the system and get the result of matching.

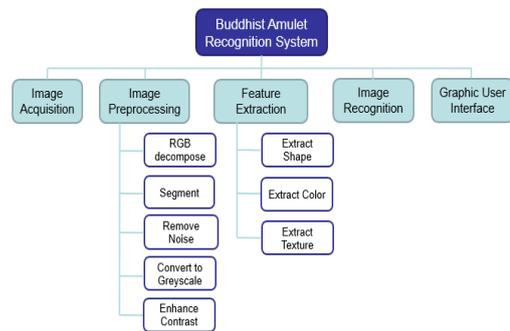


Fig. 2 Structure Chart.

From the structure chart in Fig. 2, the system is divided into five main modules, which are 1) image acquisition, 2) image preprocessing, 3) feature extraction, 4) image recognition, and 5) graphic user interface. The details of them are in the following explanation.

C. System Components

1) Image Acquisition: an amulet image has to be placed on a plain blue screen. The camera's angle is bird-eye view. It is suggested to have an amulet in the frame at least 25 percent of whole picture with complete shape. During taking the picture, flash is needed to eliminate shadow. The brightness compensation can be reduced if the image is too bright caused by flash light. The orientation of the amulet can be arranged in any angle, but its front side must be shown to the camera. The examples of the image setting for this component are shown in the Fig. 3.

2) Image Preprocessing: the image preprocessing is to prepare the amulet image before feature extraction process. This module helps the system to get important data and work more accurately. There are five sub modules in this part, which are RGB decomposition, segmentation, noise removal, grey-scale conversion, and contrast enhancement.



Fig. 3 Image setting for the system

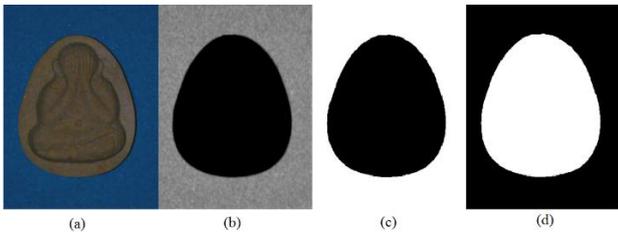


Fig. 4 Original image (a), Background indicator (b), Binary (c), Inverted binary (d)

In RGB decomposition, the input image is formed into RGB color components. They consist of three components, which are red component, green component and blue component. Each component is $m \times n$ matrices where m is image's height and n is image's width containing intensity values of their color for every pixel. The value of intensity is in an interval from 0 to 255. By this sub module, the system produces 3 matrices of red, green, and blue channel to be processed in the next module.

Segmentation is to define the object's area for proceeding in other modules. From the image acquisition module, the image background is blue, so it can be indicated that any pixels, which have intensity value in blue channel larger than red channel and green channel are pixels of background. To calculate background indicator, (1) can be used where $i =$ image's row index, $j =$ image's column index, $R =$ red value, $G =$ green value, $B =$ blue value, and $Grey =$ grey-level value.

$$Ind_{i,j} = (B_{i,j} - R_{i,j}) + (B_{i,j} - G_{i,j}) \quad (1)$$

The high value of background indicator represents background area. The low value might be the object. To justify that, it is needed to use Otsu's method [5] to find a threshold value that separate background and object. The Otsu's threshold is the value that makes the sum of background variance and foreground variance as least as possible. After Otsu's threshold are calculated, every pixel containing background indicator value higher than Otsu's threshold are mapped to be 1 or white otherwise they are 0 or black. Then, those values have to be inverted, so the object's area is turned to be white, and the background is black instead. The Fig. 4 shows an example of segmentation process.

The segmented object's area still has some noise at its perimeter, so median filter is used to remove them. It uses median value of neighbor pixels to replace the interest pixel. The result is the segmented image that have smooth object's perimeter. From this point, the white area given by this module is called as region of interest.

Grey-scale conversion is needed to analyze the texture of Buddha's image. The output from RGB the decomposition module is computerized again in this part. The output grey-scale image of this part can be produced by using (2) where $i =$ image's row index, $j =$ image's column index, $R =$ red value, $G =$ green value, $B =$ blue value, and $Grey =$ grey-level value.

$$Grey_{i,j} = 0.2989R_{i,j} + 0.5870G_{i,j} + 0.1140B_{i,j} \quad (2)$$

For contrast enhancement, the sub module is to adjust the contrast of grey-scale image derived from grey-scale conversion. To enhance gray-scale contrast, a technique

called histogram equalization is applied. The first is to create a histogram of the grey-scale image. Then it is applied with (3) where $i =$ image's row index, $j =$ image's column, $NewGrey =$ new grey value, $Grey =$ original grey value, $N =$ total number of pixels, $f_n =$ number of pixels with intensity n , $n = 1, 2, 3, \dots, 255$ [6].

$$NewGrey_{i,j} = floor \left(255 \frac{Grey_{i,j} f_n}{N} \right) \quad (3)$$

Claimed by A. Setkov et al, even the contrast is increased, it preserves the rank order of those pixels [7]. Therefore, it is invariant to brightness adjustment because the regular brightness adjustment is a monotonous color change.

In addition, the process of contrast enhancement is performed in region of interest defined by the enhanced segmented image only. The pixels outside the region are not counted.

3) Feature Extraction: this module is the process, that analyze the data given by the image preprocessing module and comes up with feature values. There are three types of focused features, which are shape, color, and texture.

In this module, the region of interest is very crucial because the background's data must not be included in the calculation.

There are two techniques needed to extract shape feature efficiently, which are centroid contour distance and Fourier transform. The combination of those techniques provides rotational and scaling invariance [8]. The centroid contour distance is the series of distance between object's centroid to all points on object's edge. The series can be plotted as a graph of signal as shown in Fig. 5.

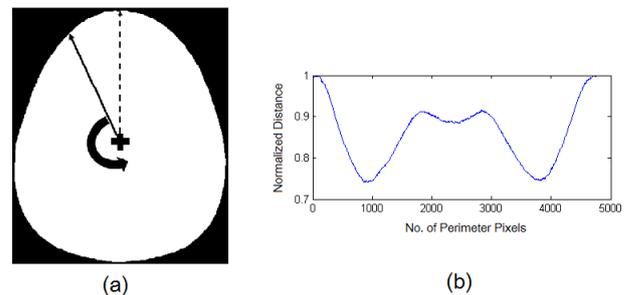


Fig. 5 Perimeter tracing (a), Centroid contour distance graph (b)

The signal can be analyzed by using Fourier transform. The technique is to transform the signal into frequency domain. The peaks in each term of frequency can be used describing the shape of the signal. The sequence of the peak values is called as Fourier descriptor (FDT). This technique produces invariant result even the signal phrase is shifted. (4) is the equation that explains the calculation of the FDT where $N =$ total number of perimeter points, $t =$ perimeter point, $d =$ centroid distance at t , $n =$ number of term $= 0, 1, 2, \dots, 30$.

$$FDT_n = \frac{\sum_{t=0}^{N-1} d_t e^{-j2\pi nt/N}}{N} \quad (4)$$

After FDT is calculated, its amplitudes are coefficient values that represent the shape of signal. Excluding term 0, we merely use first 30 terms because the more number of terms only represents the noise in the signal [9]. The normalized form of FDT is invariant to scaled signal.

In the color feature extraction module, the RGB colors in the region of interest are calculated into six features in this part, which are red-green difference, blue-red difference, green-blue difference, red standard deviation, green standard deviation, and blue standard deviation.

To explain “difference” feature more clearly, for example, red-green difference is value of red value subtracted by green value, so, if the green value is larger than red value, the red-green difference is negative.

Those color features are used because they are relative values. This is the plan to mitigate dynamic brightness compensation problem.

In the section of texture extraction, the strategy is to divide the grey-scale image in pieces, but regular dividing cannot provide rotation invariance to the system. The tactic is to use circles to divide it into ring-based regions as the Fig. 6. Using the object’s centroid as center, draw 9 circles around it where d is longest centroid contour distance of the amulet by following radii which are: $d, 0.8*d, 0.7*d, 0.6*d, 0.5*d, 0.4*d, 0.3*d, 0.2*d,$ and $0.1*d$. The circles divide the amulet into nine rings as shown in Fig.6. Each ring’s grey values are calculated to find mean and standard deviation.

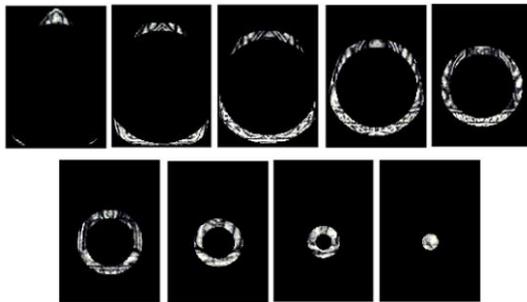


Fig. 6 Ring divisions

To sum up, there are totally 54 features derived from feature extraction process, which are 1-30 Fourier coefficient values, red-green difference, blue-red difference, green-blue difference, red standard deviation, green standard deviation, blue standard deviation, ring 1-9 mean values, and ring 1-9 standard deviation values.

4) Image Recognition: artificial neural network (ANN) is a technique for the recognition process [10]. The technique is mathematical model inspired from biological nervous system in animal brain. The network consists of many cells called “neurons” interconnected to each other with “synapses”. The neurons and synapses work together. The electric signal from one neuron stimulates other neurons via synapses. If there are more synapses between them, the neurons will be more sensitive to incoming signal.

From the concept, it comes up with the model in the Fig. 6 [11]. The input values are sent into input neural nodes as input neurons. They are multiplied by weights of each connection before sending to the output node. The weights represent how strong of synapse between input neurons and output neurons. The output node will sum those weighted values including bias and calculate them by activation function for its output.

There are many different models of ANN, but two basic models are claimed in this research, which are perceptron and multi-layer perceptron. The perceptron is an ANN that has only two layers, which are input layer and output layer. Unlike perceptron, multi-layer perceptron has more than two layers,

which are input layer, hidden layer(s), and output layer as shown in the Fig. 8.

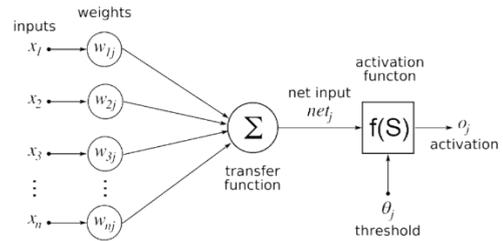


Fig. 7 How artificial neurons proceed

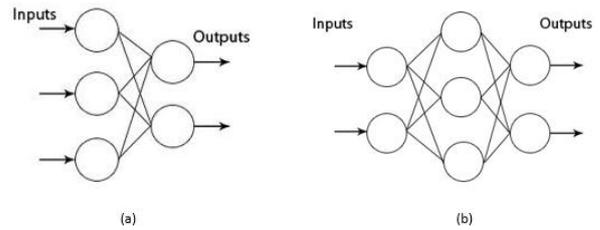


Fig. 8 Perceptron (a), Multi-layer perceptron (b)

To design the perceptron and multi-layer perceptron for this research, the number of input nodes has to be 54 because of the number of involved features. The output nodes number is referred to the number of amulet classes in the experiment’s scope which is 100. Each output nodes represent 1 amulet class from 100 classes, so the activated node is the result of the network. The result is concluded as unknown amulet if there is no or many output nodes activated. For multi-layer perceptron, the number of hidden nodes is defined as 54 which are the most possible smallest number [12], and it can be increased if the 54 is not sufficient. The activation function of perceptron and multi-layer perceptron are described in (5) and (6) [13] respectively where $x = \text{sum of weighted input}$. (5) is called as hard limiter that produce a value of 0 or 1. (6) is non-linear function called Logistic Sigmoid function. Its value is in the interval $[0, 1]$.

$$f(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (5)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

5) Graphic User Interface: from the Fig. 9, there are six parts in the system’s graphic user interface, which are 1) input image section, 2) result image section, 3) result information section, 4) feature values section, 5) ANN model section, 6) image location, 7) browse button, and 8) start button. To use the interface, the user needs to click at browse button select “7” an amulet picture in JPEG format. Then, the system will show the picture in the section of input image “1” with path of the file in image location “6” and calculate its features displayed on feature values section “4”. To select model of neural network, the user can choose by clicking radio button in ANN model section “5”. After that, the user has to click at start button “8” to initialize matching process. Finally, the result image is shown in result image section “2” with its information in result information section “3”.

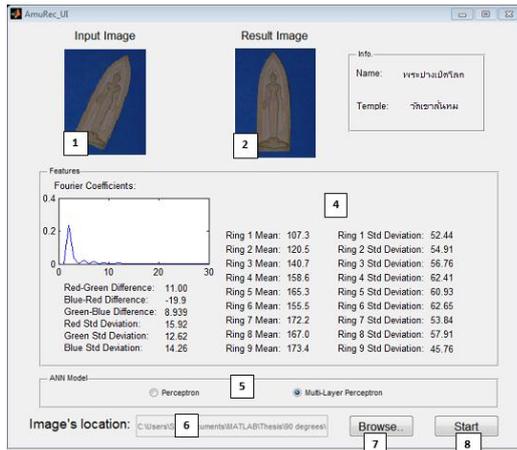


Fig. 9 Graphic User Interface

D. System Training and Testing

Because the perceptron and multi-layer perceptron are supervised learning machine, they need expected results to learn. Back-propagation training is a common method in their learning process. In each iteration, the input data for training is passed into the network. Then it comes up with outputs and errors. The errors and gradients are calculated to adjust weights depending on each connected output nodes. How much the weights changed is related to the learning factor. At the end of all iteration, it is needed to measure the network's performance by using performance function such as mean absolute error (MAE), or mean squared error (MSE). In this experiment, mean absolute error function is used for the perceptron and mean squared error for multi-layer perceptron. They might use the different performance functions because the multi-layer perceptron produces very small differences among iterations, so it uses MSE to clarify the difference by squaring its mean error. The process is running iteratively until the network reaches a stopping condition. For examples, the criteria can be to define maximum iteration or minimum MSE. In this experiment, the network is trained by 1,000 images of 100 amulet kinds (10 images per kind). For the perceptron, the stopping condition is $MAE = 0$, but the stopping point or MSE of the multi-layer perceptron is varied at 10^{-3} , 10^{-4} , and 10^{-5} to find the most efficient point.

In testing phase, the test set consists of 103 kinds of amulet – 100 images for trained or known kinds, 3 for unknown kinds, and 5 samples for each kind. Therefore, the number of images in test set is 515.

IV. EXPERIMENTAL RESULT

Firstly, based on the Table I, the perceptron is tested. It can match the trained amulets correctly for 489 from 500 images. For unknown type, the system classify them as unknown for 14 from 15 images. In total, the accuracy rate of the perceptron is 97.67 percent.

According to Table II, the multi-layer perceptron that is stopped in training at 10^{-3} of MSE can match known types for 445 from 500 images, and return unknown result for unknown types by 15 from 15 images. Totally, its accuracy rate is 89.32 percent.

Refer to Table III, the multi-layer perceptron with $MSE = 10^{-4}$ is able to show matched result for 490 from 500 images in trained data set, and classify unknown data set as unknown

type for 15 from 15 images. Its accuracy is calculated as 98.06 percent.

The last is for Table IV. The multi-layer perceptron with a stopping condition: $MSE = 10^{-5}$ can match 500 trained data and classify 15 unknown data as unknown type, so its accuracy rate is 100 percent.

TABLE I: PERCEPTRON'S RESULT

| Type of Dataset | Match | Mismatch | Unknown |
|-----------------|----------------|--------------|---------------|
| Known | 489/500 | 3/500 | 8/500 |
| Unknown | 0/15 | 1/15 | 14/15 |
| Total | 489/515 | 4/515 | 22/500 |

TABLE II: MULTI-LAYER PERCEPTRON'S RESULT ($MSE=10^{-3}$)

| Type of Dataset | Match | Mismatch | Unknown |
|-----------------|----------------|--------------|---------------|
| Known | 445/500 | 0/500 | 55/500 |
| Unknown | 0/15 | 0/15 | 15/15 |
| Total | 445/515 | 0/515 | 75/500 |

TABLE III: MULTI-LAYER PERCEPTRON'S RESULT ($MSE=10^{-4}$)

| Type of Dataset | Match | Mismatch | Unknown |
|-----------------|----------------|--------------|---------------|
| Known | 490/500 | 0/500 | 10/500 |
| Unknown | 0/15 | 0/15 | 15/15 |
| Total | 490/515 | 0/515 | 25/515 |

TABLE IV: MULTI-LAYER PERCEPTRON'S RESULT ($MSE=10^{-5}$)

| Type of Dataset | Match | Mismatch | Unknown |
|-----------------|----------------|--------------|---------------|
| Known | 500/500 | 0/500 | 0/500 |
| Unknown | 0/15 | 0/15 | 15/15 |
| Total | 500/515 | 0/515 | 15/515 |

In training phrase, the perceptron needs 207 iterations to reach $MAE = 0$. For multi-layer perceptron, it needs 810, 1693, and 3123 iterations to reach $MSE = 10^{-3}$, 10^{-4} , and 10^{-5} respectively.

In term of performance, on average, the system need 7.18 seconds per image to proceed image preprocessing and feature extraction, 0.19 second per image for perceptron recognition, and 0.18 second per image for multi-layer perception recognition.

V. CONCLUSION

To conclude, the multi-layer perceptron where $MSE = 10^{-5}$ produces the best result. It can classify the data set in this experiment perfectly. The regular perceptron also have a good result, but it still has some mistakes in known and unknown set classification.

The designed image preprocessing and feature extraction techniques are effective under the defined environment of this research.

The research objective is fulfilled, and it is expected to be expanded in the future.

REFERENCES

- [1] P. Thumwarin, S. Malila, P. Janthawong, and W. Pibulwej, "A Robust Coin Recognition Method with Rotation Invariance," in *Proc. International Conference on Communications, Circuits and Systems*, Guilin, 2006, pp. 520-523.
- [2] C. Korbuakaew, "Identification of Amulets with Special Feature Matching," M.S. thesis, Dept. Computing., Silpakorn Univ., Bangkok, Thailand, 2007.
- [3] J. Wongkorsub, T. Pornaudomdaj, P. Vessawadsi, and C. Pornpanomchai, "Buddhist Amulet Recognition System (BARS),"

in *Proc. 2nd International Conference on Computer and Network Technology*, Bangkok, 2010, pp. 495-499.

- [4] N. Srisuporn and C. Pornpanomchai, "Buddhist amulet coin recognition by genetic algorithm," in *Proc. 17th International Computer Science and Engineering Conference*, Nakorn Pathom, 2013, pp. 324-327.
- [5] A. Greensted. (Jun. 17, 2010). *Otsu Thresholding* [Online]. Available: <http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.htm>
- [6] R. L. Eastern Jr. (Nov. 22, 2010). *Fundamental of Digital Image Processing* [Online]. Available: https://www.cis.rit.edu/class/simg361/Notes_11222010.pdf
- [7] A. Setkov, M. Gouiffes, and C. Jacquemin, "Color Invariant Matching for Image Geometric Correction," in *Proc. 6th International Conference on Computer Vision / Computer Graphics Collaboration Techniques and Applications*, 2013©ACM. doi: 10.1145/2466715.2466731
- [8] D. Zhang and G. Lu, "A Comparative Study on Shape Retrieval Using Fourier Descriptors," in *Proc. 5th Asian Conference on Computer Vision*, Melbourne, 2002, pp. 646-651.
- [9] Z. S. Tseng. (2012). *Second Order Linear Partial Differential Equation* [Online]. Available: <http://www.math.psu.edu/tseng/class/Math251/Notes-PDE%20pt2.pdf>
- [10] J. Heaton, *Introduction to Neural Networks with Java*, 2nd ed. Missouri: Heaton Research, 2008, ch. 1, pp. 47-48.
- [11] WikiBooks. (Mar. 14, 2012). *Artificial Neural Networks* [Online]. Available: http://en.wikibooks.org/wiki/Artificial_Neural_Networks/Print_Version
- [12] J. Heaton, *Introduction to Neural Networks with Java*, 2nd ed. Missouri: Heaton Research, 2008, ch. 5, pp. 158-159.
- [13] J. Heaton, *Introduction to Neural Networks with Java*, 2nd ed. Missouri: Heaton Research, 2008, ch. 5, pp. 151.



Waranat Kitiyanan was born in Bangkok Thailand. He studied in Bachelor of Science programme of the faculty of Information and Communication Technology, Mahidol University and graduated in 2012.

He is studying in computer science programme of the same faculty for the master degree in science.



Chomtip Pornpanomchai received his B.S. in general science from Kasetsart University, M.S. in computer science from Chulalongkorn University and Ph.D. in computer science from Asian Institute of Technology. He is currently an assistant professor in the faculty of Information and Communication Technology, Mahidol University, Bangkok, Thailand. His research interests include artificial intelligence and pattern recognition.