Automatic Identification of Human Helminth Eggs on Microscopic Fecal Specimens Using Digital Image Processing and an Artificial Neural Network

Yoon Seok Yang, *Student Member, IEEE*, Duck Kun Park, Hee Chan Kim*, *Member, IEEE*, Min-Ho Choi, and Jong-Yil Chai

***Abstract*—In order to automate routine fecal examination for parasitic diseases, we propose in this study a computer processing algorithm using digital image processing techniques and an artificial neural network (ANN) classifier. The morphometric characteristics of eggs of human parasites in fecal specimens were extracted from microscopic images through digital image processing. An ANN then identified the parasite species based on those characteristics. We selected four morphometric features based on three morphological characteristics representing shape, shell smoothness, and size. A total of 82 microscopic images containing seven common human helminth eggs were used. The first stage (ANN-1) of the proposed ANN classification system isolated eggs from confusing artifacts. The second stage (ANN-2) classified eggs by species. The performance of ANN was evaluated by the tenfold cross-validation method to obviate the dependency on the selection of training samples. Cross-validation results showed 86.1% average correct classification ratio for ANN-1 and 90.3% for ANN-2 with small variances of 46.0 and 39.0, respectively. The algorithm developed will be an essential part of a completely automated fecal examination system.**

***Index Terms*—Artificial neural network, automatic identification, helminth egg, pattern classification.**

I. Introduction

HUMAN parasites are classified taxonomically into protozoa and helminths (nematodes, trematodes, and cestodes). Parasites enter the human body mainly through soil, water, vegetables, fruits, or other host animals, then inhabit internal organs such as the small intestine, liver, and lung. The infected human body suffers from malnutrition, irritation, damaged tissue, parasitic contamination, greater likelihood of microorganizm intrusion, allergic response by the body itself, and in the worst case, death.

In developed countries, parasitoses are less prevalent, with fewer people becoming ill from parasitic diseases. However, the extended global transportation network, which has encouraged more overseas travel and international food fads, allows parasitic diseases to migrate from one country to other distant countries. And despite high levels of sanitation, food inspection, and vector control in developed countries, individuals who may have acquired parasitic infections elsewhere require medical treatment. In addition, the role of zoonotic parasite is becoming more important as agents of human diseases.

Human parasitic infections are recognized through the identification of parasitic organizms in the feces, urine, blood, and tissues by using the proper diagnostic methods [1]. Fecal examination is the most common diagnostic method. Trained experts examine the fecal specimens, searching for parasitic organizms, usually the eggs of helminths and cysts of protozoa. If the organizms are present, they examine the sizes, shapes, and numbers to identify the species of parasites, the degree of infection, and appropriate therapeutic modalities. But diagnosticians are reluctant to perform fecal examinations; with decreasing importance of most parasites, less material is submitted for examination, which results in an insufficiency of training material. Thus, it is difficult to maintain enough people with expertise in diagnostic parasitology. Worst of all, diagnostic parasitologists are expected to maintain high standards of proficiency in the identification of rare parasites as well as common endemic ones.

To resolve the issue of a lack of trained personnel and relieve the pressure on them, the automation of routine fecal examinations is necessary. The complete automation of check-up routines would allow inspection of a great number of specimens with a high degree of reliability and objectivity. This would be useful in developed countries as well as in underdeveloped and developing countries.

Several precedent studies employed microscopic image analysis to diagnose parasites [2]–[6]. These studies successfully differentiated several parasites in pigs and cattle with numerically extracted morphologies through the use of interactive image analysis tools [2], [3], [6]. They extracted morphometric features, including length, smoothness, and sphericity with subjectively chosen parameters from digital microscopic images [2], [3]. These features were subsequently involved in the computation of classification indices; the parasite species were then identified using the calculated indices. Their works were concentrated on statistical analysis of morphologic values and round-robin discrimination tests. However, no truly automatic and computer-assisted algorithm has been developed. Specifically, this would entail automation of the entire procedure, including image enhancement, extrac-

Manuscript received January 4, 2000; revised July 2, 2001. This work was supported in part by the Year 2000 BK 21 Project for Medicine, Dentistry and Pharmacy. *Asterisk indicates corresponding author*.

Y. S. Yang and D. K. Park are with the Interdisciplinary Program Biomedical Engineering Major, Seoul National University, Seoul 110-744, Korea.

*H. C. Kim is with the Departments of Biomedical Engineering, Seoul National University College of Medicine and Institute of Biomedical Engineering, Medical Research Center, Seoul National University, 28 Yongon Dong, Chongno Gu, Seoul 110-744, Korea (e-mail: hckim@snu.ac.kr).

M.-H. Choi and J.-Y. Chai are with the Department of Parasitology, Seoul National University College of Medicine and Institute of Endemic Diseases, Seoul National University Medical Research Center, Seoul 110-744, Korea.

Publisher Item Identifier S 0018-9294(01)04170-2.

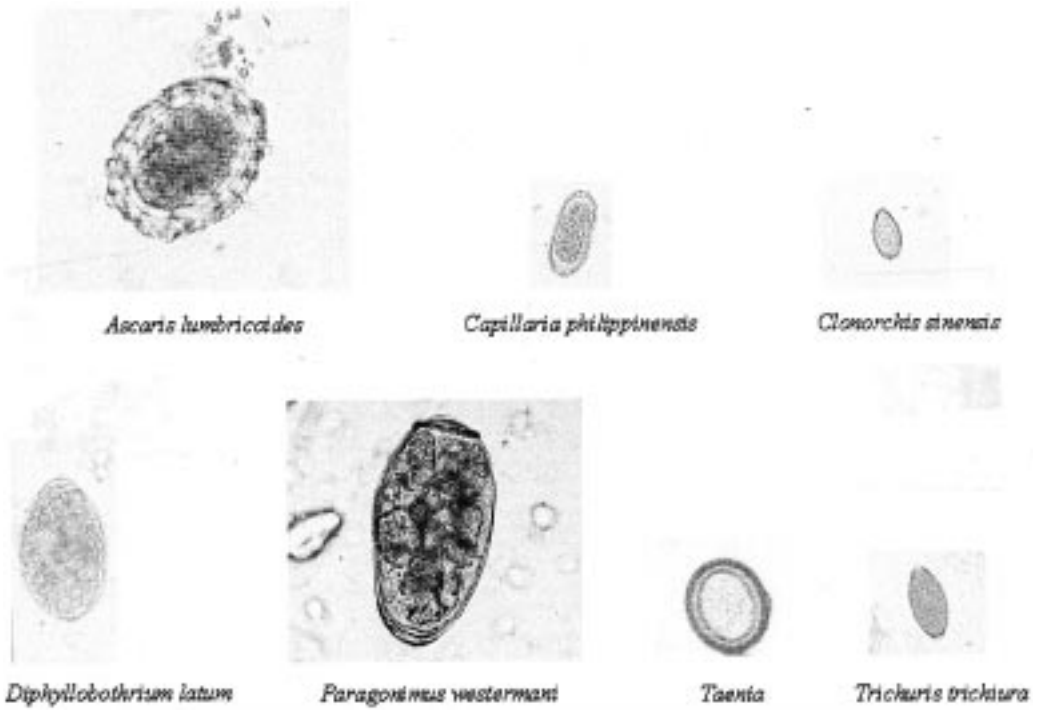


Fig. 1. Representatives of the seven helminth eggs.

tion of morphological characteristics, isolation of parasitic organizms from unwanted dreg materials, and identification of parasite species. To develop a totally automated system, image segmentation, the algorithm for morphometric feature extraction, and classification should be automated. In addition, dedicated hardware equipment is necessary to scan complete images of each specimen under microscope.

We propose in this paper an automated method to detect common helminth eggs in microscopic fecal specimen images and to identify their correct species using digital image processing techniques [7] and an artificial neural network (ANN) system [8]. The ANN classifier is widely used in pattern recognition problems owing to its good performance and easy application [8]–[10].

II. Materials and Methods

Trained laboratory staff normally examines fecal specimen visually under 100-fold magnification. Although the higher magnification views are occasionally used to examine cell material inside the egg shell, it is sufficient for pinpointing a parasitic egg on the specimen and checking its size and shape. A relatively faster scanning speed is one advantage over the higher magnification. For digital processing, microscopic images were digitized by using a color video camera (JVC, Japan) embedded in the microscope (Olympus, Japan) and an image capturing board (MERIT, Korea) on a Pentium II PC. Images were captured under 200-fold magnification to provide comparable quality with a natural view under 100-fold magnification. Each digital image is 560 by 420 pixels large with 8-bit intensity depth, covering a slide area of 130 μm $\times$100 μm. This consequently yields a spatial resolution of 4.2 pixels/μm. A total of 82 images were acquired from formalin-preserved specimens containing eggs of seven human helminths. There were about 12 images/parasite species. The seven typical species of human helminth were chosen as follows [1]: *ascaris lumbricoides, trichuris trichiura, capillaria philippinensis, clonorchis sinensis, paragonimus westermani, diphyllobothrium latum, and taenia*. Representatives of each of the egg types are shown in Fig. 1 to help the reader understand the classification problem visually.

To localize and classify helminth eggs in a given image, we first found meaningful objects in the image and extracted their features using digital image processing methods. Then a two-stage ANN classified them based on their numeric features. All processing routines were written in MATLAB (Mathworks, Natick, MA) and overall processing steps are outlined in Fig. 2.

A. Digital Image Processing

Digital image processing consists of four operational steps: median filtering, binary thresholding, segmentation, and feature extraction.

1) Median Filtering: Since fecal material normally includes blood residue, various products of digestion, sloughed epithelial cells, mucus, and microorganizms such as bacteria and yeast, the digitized images show nonhomogeneous background characteristics as in Fig. 3 [1]. Therefore, all images were preprocessed with a 4×4 median filter to remove small unimportant dregs without degrading overall image quality. Filter dimensions were experimentally determined to remove small particles while preserving target objects.

2) Binary Thresholding: A relatively simple binary thresholding method was applied to the median filtered image. Compared with other edge detection methods that yield problems with cracked boundary lines, it was proven to have an acceptable

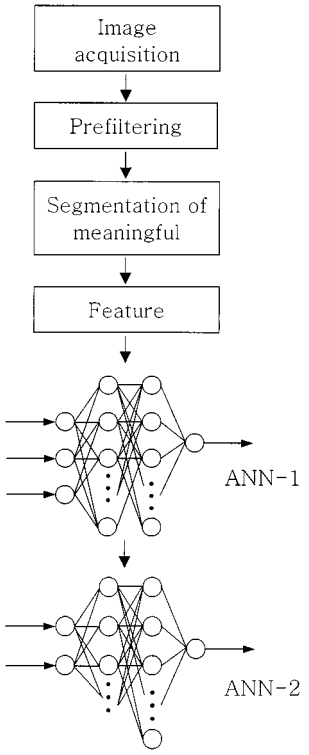


Fig. 2. Overall processing steps.

performance with minimal complexity in this specific application. The threshold was empirically set at 35% of the highest intensity level in all cases. It was determined based on the fact that most of the histograms have bimodal distribution and their valley points coincide with this level within reasonable range. It is well known that the valley point of the bimodal distribution ensures the optimal classification result in terms of Bayesian decision errors [10], [11].

3) Segmentation: From the thresholded binary results, we segmented only objects whose sizes were above the minimum of egg sizes (600-pixel dimension) by the connected component operator [7]. Binary thresholding may have caused the segmented objects not to appear as one full mass. In order to make later numerical feature extraction easier, every segmented object was filled by the four-directional inward region growing algorithm so that there were no internal holes. Then the single skeleton algorithm produced a boundary line of an object with one-pixel thickness for truly quantitative representations of the object's morphology [7].

4) Feature Extraction: As criteria for detection and identification of helminth eggs, three morphological characteristics were selected as follows [1].

a) *Size*: Helminth eggs ranged in length from 25 μm to 150 μm, and in diameter from 12–14 μm to 90 μm. Typical sizes of common helminth eggs are compared in Fig. 4.
b) *Shape*: Each parasite species produces eggs that are uniform in shape. The shape of eggs, varying from spherical to elongated, is a prominent feature for most species of helminths.
c) *Eggshell*: Helminth eggs usually have a smooth shell. With the exception of *A. lumbricoides*, these eggs have a mammillated external shell layer, the smoothness of the shell is an important feature in separating eggs from vegetable cells or other plant material that may have irregularly shaped membranes.

To obtain numeric values representing each characteristic, we calculated the polar coordinates (r, θ) of all detected boundary pixels with the origin at their gravitational center. An one object's boundary line can then be represented by the trajectory of its radial distance, $r(\theta)$, as θ spans from zero to 2π, as illustrated in Fig. 5 [7]. After normalization of $r(\theta)$ by its mean value, three numeric features representing shape and shell characteristics are calculated. The object's size in total pixels number is included as the fourth numeric feature. All four features are summarized as follows.

a) $F_1 = \min\{r(\theta)\}/\max\{r(\theta)\}$, the ratio of minimum in $r(\theta)$ to maximum in $r(\theta)$.
b) $F_2 = \sum_0^{\pi/4}|R(\omega)|/\sum_{\pi/4}^{\pi}|R(\omega)|$, the ratio of low frequency components to high frequency components from the magnitude of discrete Fourier transform of $r(\theta)$, $|R(\omega)|$,
c) $F_3 = \sigma_r$, standard deviation of $r(\theta)$.
d) F_4, the number of pixels in an object area or pixel dimension.

B. ANN Classifier

Species identification was performed by an ANN classification system which consists of two stages; the first stage (ANN-1) separates parasite eggs from the artifact, while the second stage (ANN-2) identifies the species of the detected eggs. This two-stage classification scheme is adopted to reduce the total computing time that is one of the critical parameters in realizing a fully automated system. The goal of ANN-1 is to examine all segmented objects and isolate parasitic organizms from other elements in feces with similar sizes. The following ANN-2 classifies the species of the detected eggs by ANN-1 using the same numeric features. Both ANNs are general multilayer perceptrons (MLP) neural networks; the back propagation learning rule was used in both cases [8]. Generally, the MLP network consists of set of neurons (perceptrons) in several layers and inter neuronal connections. Input neurons take input values and each neuron in subsequent layers takes a sum of neuronal output activities in the previous layers weighted by their interconnections and gives its nonlinear output activity. It is basic and the most widely used neural network in pattern classification because of its capability of making almost any kind of linear or nonlinear decision boundaries between clustered patterns. MLP network capacity is limited by the number of neurons and their inter connections and by the nonlinear activation functions of these neurons.

Network topology was systematically determined in terms of variance of the classification results. For all devisable network topologies, classification results were calculated for different data sets selected by the cross-validation method. The number of layers and neurons were chosen to provide minimum variance and reasonably high average classification results (>85%). Typical results are summarized in Tables I and II for ANN-1 and ANN-2, respectively.

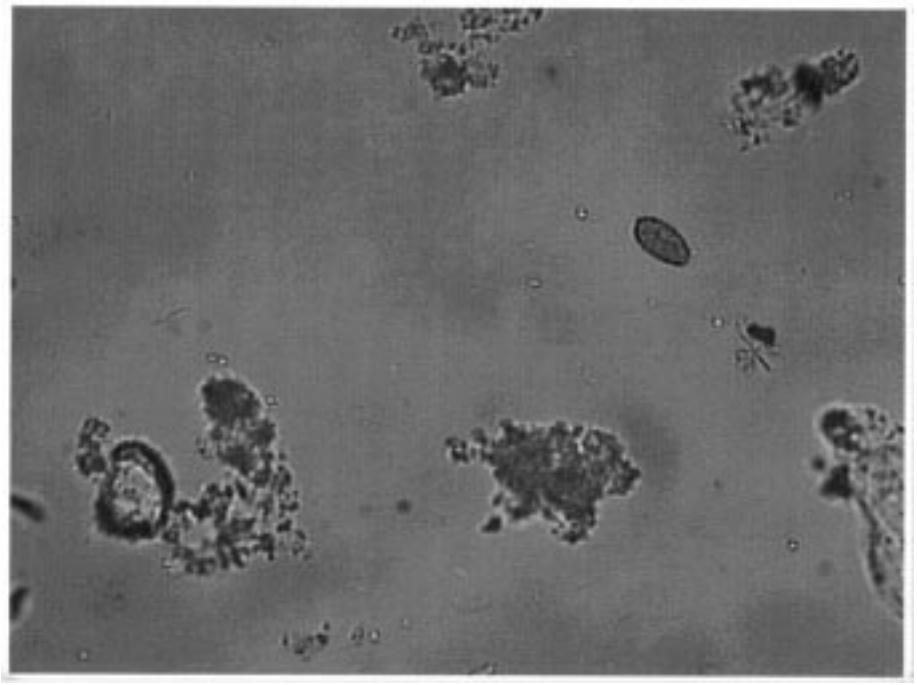


Fig. 3. One of the typical microscopic images of fecal specimen under 200-fold magnification.

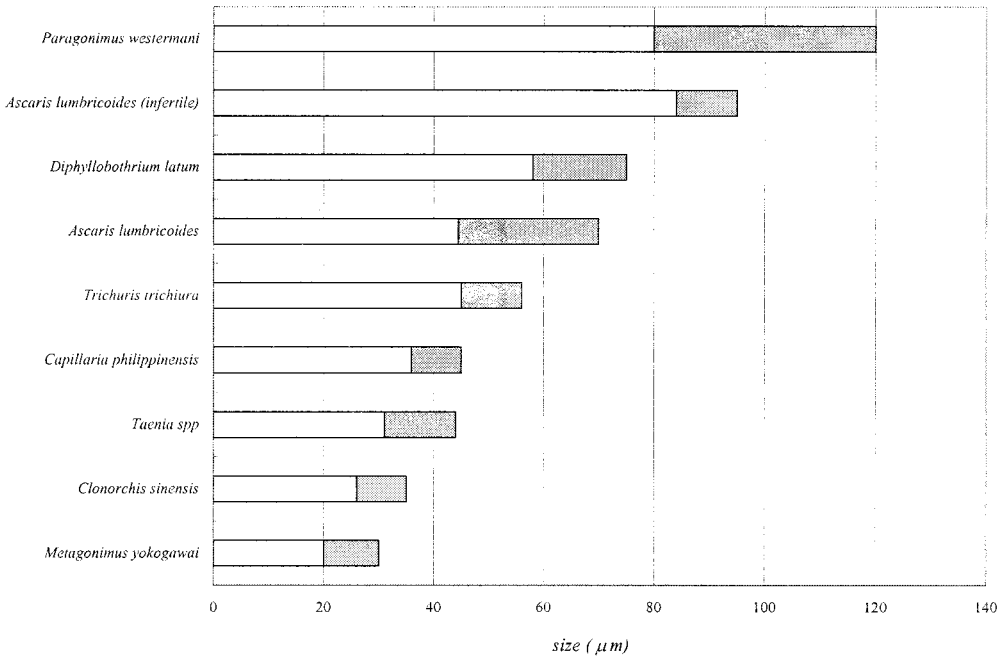


Fig. 4. Comparative sizes of typical helminth eggs. Dark bars denote ranges of length.

The cross validation took up the leave-one-out method [5] with the alteration of tenfold crossover to avoid excessive time-consuming repetitions. It generally provides network performance evaluations independent of the choice of training and test set.

1) Artifacts Removal with ANN-1: Significant morphological differences exist between parasite eggs and artifacts, even though they are similar in size, as mentioned earlier. The egg cell has a shell membrane which makes its boundary look more circular and smoother than those of the artifacts. The circularity of shell may be expressed by the ratio of its short axis and long axis. The smoothness may be rephrased as the trajectory of its radial distance, $r(\theta)$, which has smaller high frequency components than those of the artifacts. The input layer has four input nodes to receive the morphological feature values. It has one output node which indicates whether the input morphologies are those of eggs or artifacts. The desired output values were derived from the ground truth obtained from diagnostic experts. A standard sigmoid activation function was used at the input and hidden nodes, while a linear activation function was used at the output node [8]. 187 samples (87 eggs and 100 artifacts in total) were randomly divided into ten subsets of approximately equal

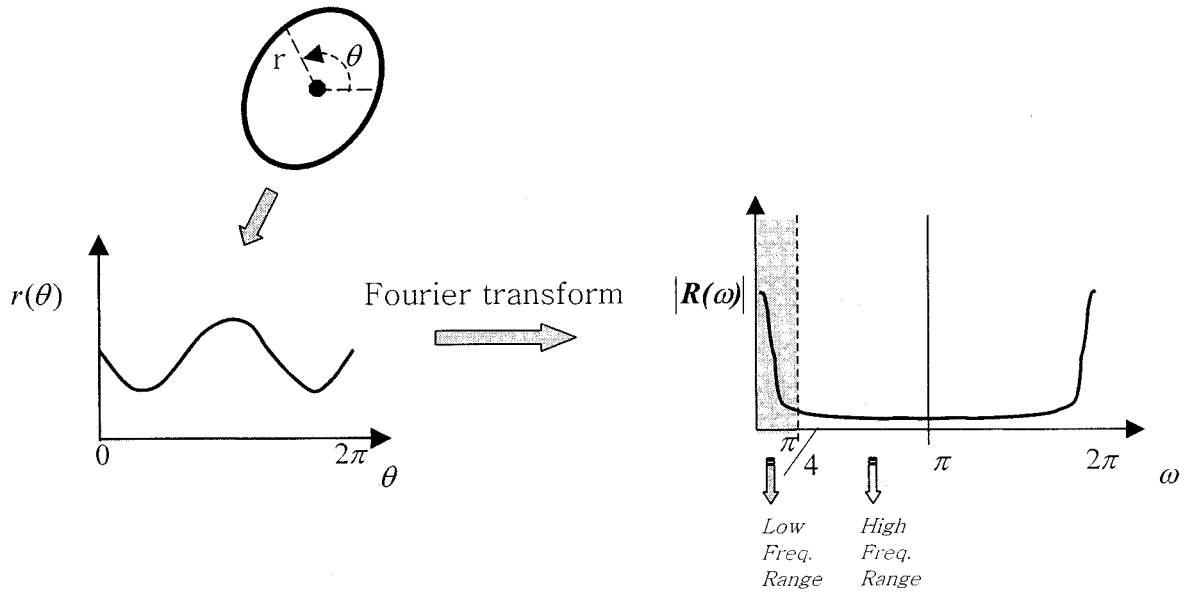


Fig. 5. Schematic diagram of calculating angular trajectory $r(\theta)$ from boundary line and its frequency components $R(\omega)$. $r(\theta) \xrightarrow{F} R(\omega)$ where F denotes the Fourier Transform.

TABLE I
CROSS-VALIDATION RESULTS OF ANN-1 FOR DIFFERENT NETWORK TOPOLOGIES

Number of neurons in		Correct classification ratio (%)	
1st hidden layer	2nd hidden layer	Average	Variance
20	0	84.5	102.6
8	8	86.1	46.0
14	14	85.6	80.3

TABLE II
CROSS-VALIDATION RESULTS OF ANN-2 FOR DIFFERENT NETWORK TOPOLOGIES

Number of neurons in		Correct classification ratio (%)	
1st hidden layer	2nd hidden layer	Average	Variance
12	0	87.3	69.0
8	8	90.0	39.0
14	14	87.0	44.0

size ($187/10 \cong 19$) for tenfold cross validation. After systematic evaluation, an ANN topology having eight neurons in the first hidden layer and eight neurons in the second hidden layer was chosen since it shows the maximum average and the minimum variance of classification results.

With this topology, the neural network was trained with randomly selected halves of 87 eggs and 100 artifacts as training samples. Then classification results with the remaining data samples were computed. Besides, since the results can be affected by cutoff at the output neuron, the receiver operating characteristic (ROC) was also analyzed to show the efficiency in true positive (TP)/false positive (FP) tradeoff [11].

Among segmented objects in images, those which yielded outputs above cutoff value at the trained ANN-1 were determined to be eggs and transferred to the next identification stage.

2) Identification of Helminth Species: The ANN-2 has three input nodes and seven output nodes. The three input nodes receive the same numeric morphological feature values that are used in the ANN-1 except for the third element. A preparatory classifier using all 4 features including the third element (σ_r) produced no significant improvement over the three input classifier for all network topologies. A reasonable explanation for this is that most of the eggs, once distinguished from various artifacts, have σ_rs within a narrow range. The computed variance of the σ_r over all detected eggs was smaller than those of other three features. The features used in ANN-1 and ANN-2 are summarized in Table III for clarity. Contributions of individual features to the ANN-2 classification are described by the following examples. F_4, the size of the egg, is the strongest characteristic in finding *C. sinensis*, which is remarkably smaller than the other kinds of eggs, as shown in Fig. 4. The smoothness of shell is well represented by F_2; this is important especially in distinguishing *A. lumbricoides* from *D. latum*. Despite their similarity in size, they are quite different in terms of F_2 since the former, especially in unfertilized cases, has a rougher shell than the latter. Finally, F_1, which corresponds to the ratio of the short axis to the long axis of the oval boundary helps to separate circular eggs of *Taenia* from the more elliptical eggs of *T. trichiura* and *C. philippinensis*.

The seven output nodes correspond to the seven helminth species selected for this study. ANN-2 has a competitive output layer in which only one node produces a valid output value for each input pattern. Such "winner-take-all" competition is suitable for this type of classification problem [8]–[10]. Each artificial neuron's activation function and training rules were the same that used for ANN-1. 87 samples (87 eggs in total) were divided into ten subsets of approximately equal size ($87/10 \cong 9$) for cross validation. After systematic evaluation, the network

TABLE III
SUMMARY OF FEATURES USED IN ANN-1 AND ANN-2

Feature no.	ANN-1	ANN-2
1	$F_1 = \min\{r(\theta)\}/\max\{r(\theta)\}$	$F_1 = \min\{r(\theta)\}/\max\{r(\theta)\}$
2	$F_2 = \sum_{0}^{\pi/4}\lvert R(\omega)\rvert \Big/ \sum_{\pi/4}^{\pi}\lvert R(\omega)\rvert$	$F_2 = \sum_{0}^{\pi/4}\lvert R(\omega)\rvert \Big/ \sum_{\pi/4}^{\pi}\lvert R(\omega)\rvert$
3	$F_3 = \sigma_r$	not used
4	*Pixel dimension*	*Pixel dimension*

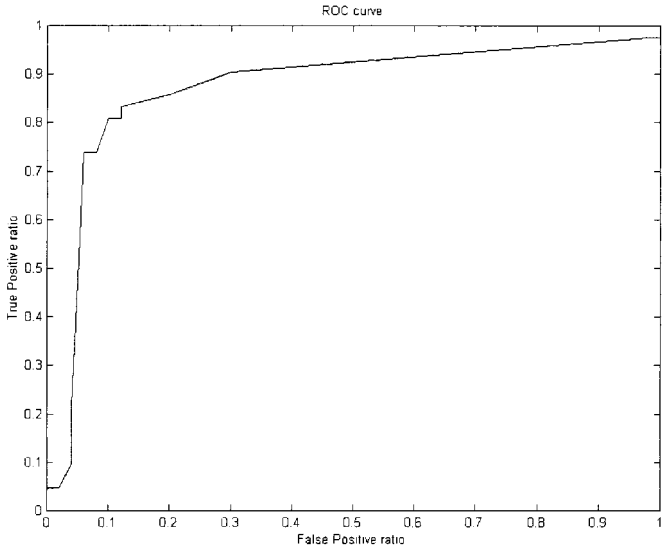


Fig. 6. ROC analysis for ANN-1. The normalized area under the curve is 0.87.

having two hidden layers with eight neurons in both layers was chosen since it showed the maximum average and the minimum variance of classification results.

Then, as in the ANN-1 case, the ANN-2 with the chosen topology was trained with randomly selected half of eggs (87 in total) and tested with the remaining half.

The neural network classifier was compared with the simple k-nearest neighbor classifier for performance. The k-nearest neighbor classifier assigns an unknown test sample to the most prevailing class among its nearest k neighbors in feature space. All the neighbors are preclassified samples given to it prior to the test [11]. The performance of k-nearest neighbor classifier was evaluated by the same cross-validation method to be compared with the ANN classifier.

III. RESULTS

A. Segmentation and Feature Extraction

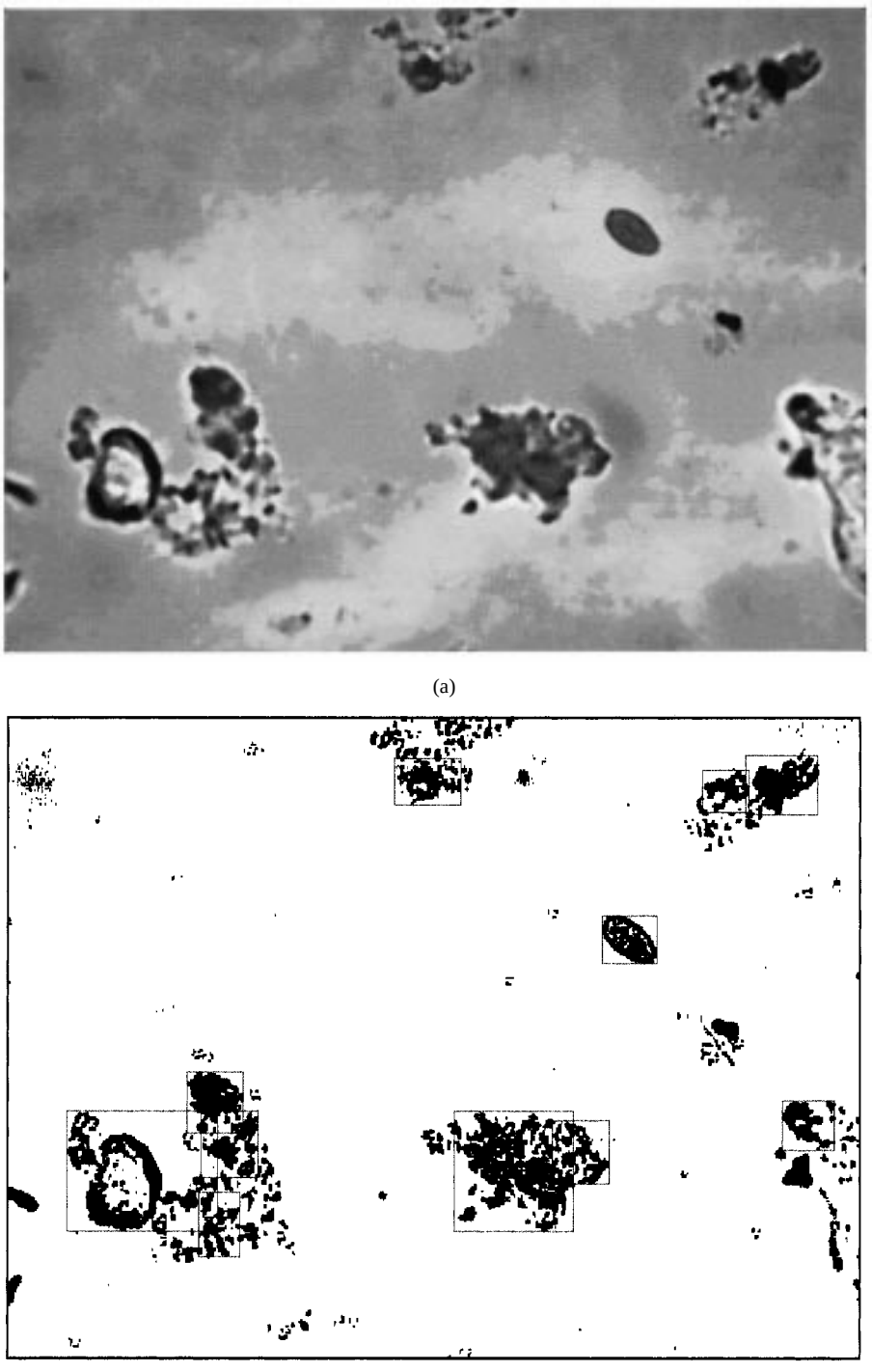
Typical digital processing results with the original image of Fig. 3 are shown in Figs. 7 and 8. In this case, the original image has one helminth egg from *C. sinensis* in the middle of the top-right quadrant. A blurred image in Fig. 7(a) is the median-filtered result of the original. It contains reduced speckles and dregs in background compared to the original. Therefore, the thresholded result in Fig. 7(b) has a suitable form for object segmentation. In this case, a total of 11 objects were selected and segmented automatically, as explained previously. They are marked on the image by square boxes. Other black spots remaining unmarked were discarded because their sizes were smaller than the minimum egg size. The segmented objects include one parasite egg and ten artifacts. In all, averaged over all images, about 15 objects were segmented and one parasite egg appeared in each image.

Fig. 8 displays four of the segmented objects consisting of three artifacts (a), (b), (c) and one parasite egg (d). It shows each object's appearance, obtained boundary line, normalized radial distance trajectory, $r(\theta)$, and log amplitude of the Fourier transform (calculated by fast Fourier transform), $R(\omega)$. The boundary line is actually the outermost pixels of filled-up version of the object and shows no seriously damaged part as expected. The radial trajectory in (d) shows a typical form of smooth elliptical egg group to which this egg belongs. The circular eggs showed more uniform distance trajectories with smaller variations than this. The other trajectories from artifacts also reflect their respective shape characteristics. The log-amplitude plots of the Fourier transform from artifacts and eggs exhibit noticeable differences. The amplitude spectrums show that the eggs contain negligible high frequency components in contrast to the artifacts.

In Fig. 9, which shows the ANN results for two different cases of eggs, numbers at the right side of each rectangle in Fig. 9(a) and (c) represents the numeric morphologic features, F_1, F_2, F_3, and F_4 extracted from $r(\theta)$ and $R(\omega)$ of the object. The values are arranged in that order.

B. ANN Classification

In all, the finally trained ANN-1 showed 100% correct classification in the training phase and 84% correct classification in the test phase. In terms of detection accuracy, it showed 83% TP ratio and 10% FP ratio on the test samples with the cutoff at the half of output scale. The results are summarized in Table IV.



(a)

(b)

Fig. 7. Typical results of digital image processing: (a) median filtered image, (b) black and white image obtained from binary thresholding. A total of 11 objects marked by rectangular boxes are detected as meaningful objects. The rectangles represent the segmented objects.

And the ROC curve in Fig. 6 shows the efficiency in TP/FP relationship. The normalized area under the curve is 0.87.

The finally trained ANN-2 showed 100% correct classification in learning stage and 83% correct classification at test stage. For more details, identification results for individual species are presented in a confusion matrix form in Table V.

In Table VI, the comparison between performance evaluations of these two ANNs and k-nearest neighbor classifiers are summarized. The three-nearest neighbor and eight-nearest neighbor classifier are compared with ANN-1 and ANN-2, respectively. The number of three and eight were chosen since they showed better performances than other k values in each application. The results show the superiority of the proposed ANN classifiers over the k-nearest neighbor classifiers.

Fig. 9 shows the typical results of final classification for the *C. sinensis* and *C. philippinensis*. The white boxes in Fig. 9(a) and (c) denotes the detected parasite eggs by ANN-1. Then the

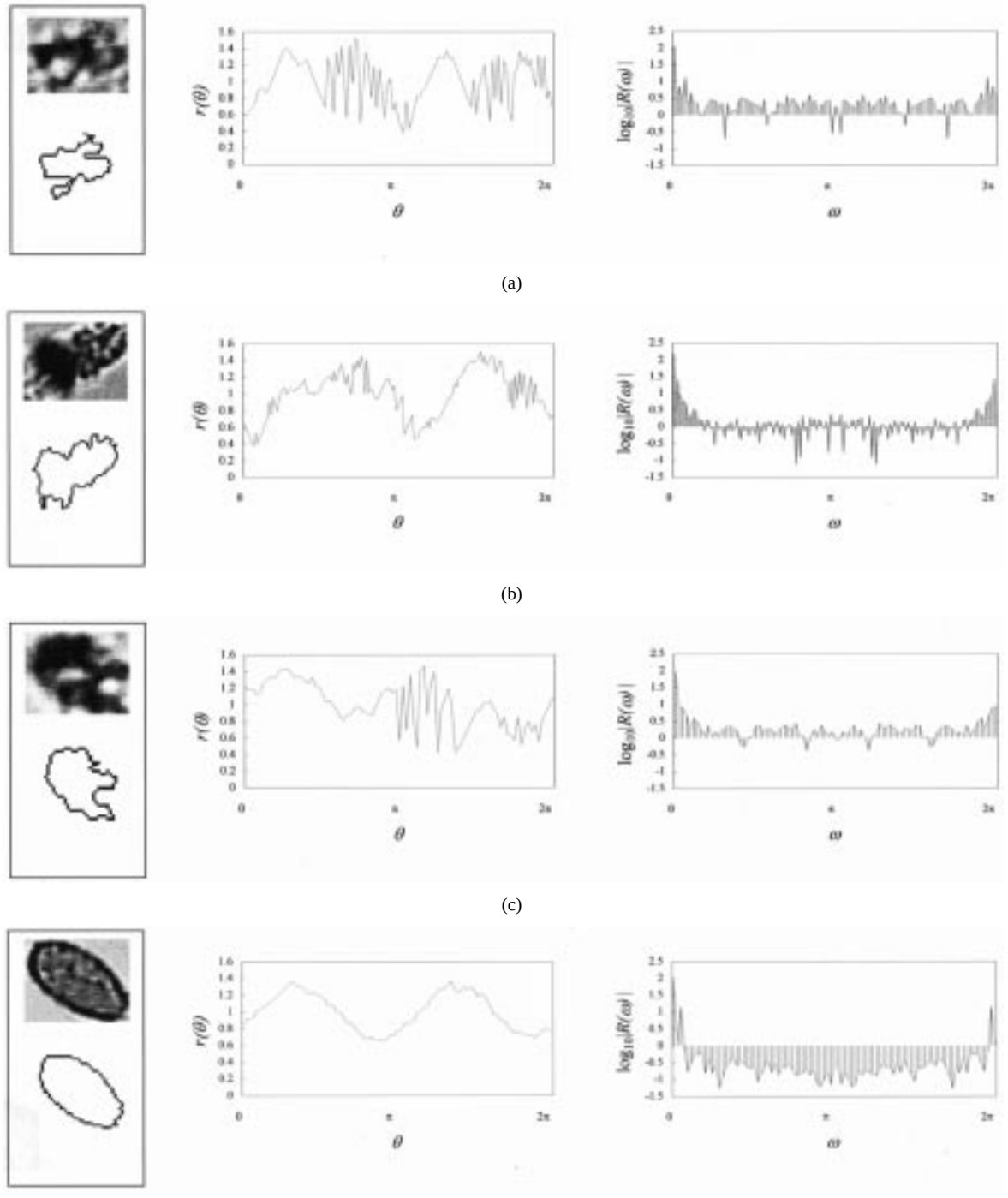


Fig. 8. Comparison of meaningful objects, segmented from Fig. 7. The leftmost images show objects and their obtained boundary lines. The plots in the middle column represent their normalized radial trajectories, $r(\theta)$. The plots in the last column show the logarithms of amplitude spectra against discrete frequencies ranging from zero to 2π. (a), (b), and (c) Artifacts. (d) Egg cell.

TABLE IV
RESULTS OF THE ANN-1 CLASSIFIER TRAINED WITH HALF OF TOTAL SAMPLES

	TP ratio (%)	FP ratio (%)
Training	100	0
Test	84	11

ANN-2 identified their species and displayed their names in Fig. 9(b) and (d), respectively.

Total computation time involving the digital image processing and ANN classifications was around 15 s/image.

IV. DISCUSSIONS

An intelligent computer program to identify helminth eggs from fecal microscopic images has been developed and verified through sample images. The developed program consists of two major steps: digital image processing and a two-stage

TABLE V
DETAILS OF ANN-2 CLASSIFICATION RESULTS IN A CONFUSION MATRIX FORM

Results \ Truth	*A. lumbricoides*	*C. philippinensis*	*C. sinensis*	*D. latum*	*P. westermani*	*Taenia*	*T. trichiura*
A. lumbricoides	2/3						
C. philippinensis		7/7					3/4
C. sinensis	1/3		6/8				
D. latum			1/8	5/6			
P. westermani				1/6	6/6		
Taenia						8/8	
T. trichiura			1/8				1/4
Total 83%	***2/3***	***7/7***	***6/8***	***5/6***	***6/6***	***8/8***	***1/4***

TABLE VI
COMPARISONS OF ANN-1 AND ANN-2 WITH SIMPLE k-NEAREST NEIGHBOR CLASSIFIERS

Classifiers		Correct classification ratio (%) in leave-out test	
		Average	Variance
Detection	ANN–1 results	86.1	46.0
	3-Nearest neighbor	74.5	75.7
Identification	ANN-2 results	90.0	39.0
	8-Nearest neighbor	75.7	76.5

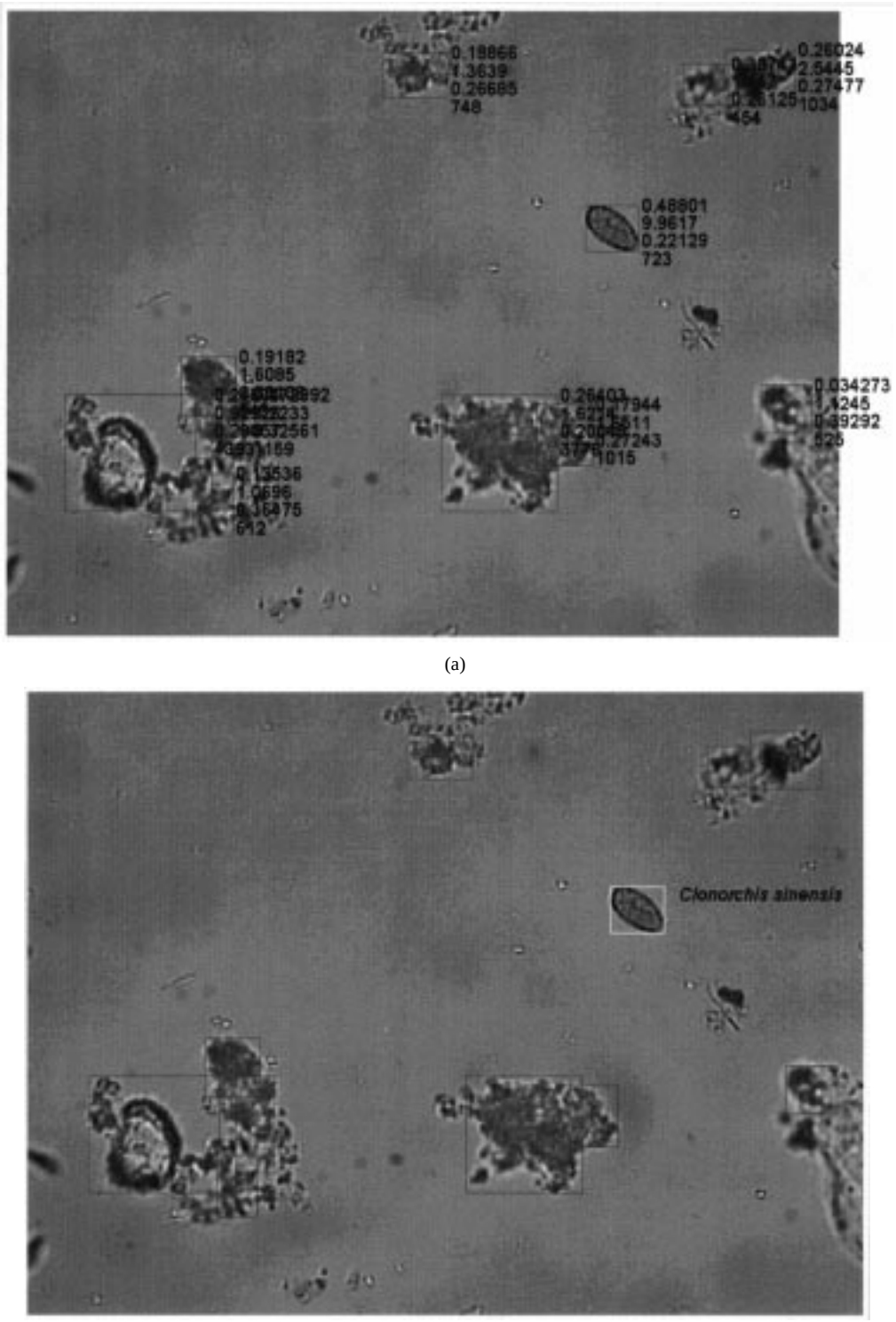
ANN-based classifier. The results showed acceptable performances with an 84% detection ratio and an 83% identification ratio and proved the applicability of the developed algorithms to full automatic examination system.

The visual recognition results used as ground truths in this study are considered very reliable since they had been derived from well-trained human experts and also cross checked by other diagnostic methods, for example, DNA analysis.

The ANN errors revealed in the detection and identification results are thought to be caused mainly by incomplete segmentation results rather than by the ANN performance itself. The errors involved in poorly segmented objects affect ANN at the training phase as well as the test (application) phase by providing inaccurate numeric features. Despite its advantages over the common edge detection methods specifically for microscopic images, the binary thresholding method used in this study results in tradeoff between accuracy and complexity. In majority of images, this method showed good segregation result. However, it produced erroneous results in minor cases especially when there is any nearby material around the object of concerns. More complicated and accurate segmentation techniques such as [12]–[17] are expected to improve the results and being investigated for possible simpler form.

The practical fecal examination is based not only on the existence of helminth eggs but also on the number of their findings. Since the helminth normally produces a large number of eggs, there must be many of them in a fecal specimen. Therefore, any few positive findings, whether True or False, need further examinations such as blood and tissue analysis.

There is no general rule for selection of proper network topology and proper data sample size for the best performance. We studied the dependency of the optimal network topology on the available data size by comparing the results with available data of 52 and 82. The numbers of neurons (input–1st hidden–2nd hidden–output layer) in ANN-1 and ANN-2 for the best results were changed from 4-6-6-1 to 4-8-8-1 and from 3-6-6-7 to 3-8-8-7, respectively. This indicates that there exists



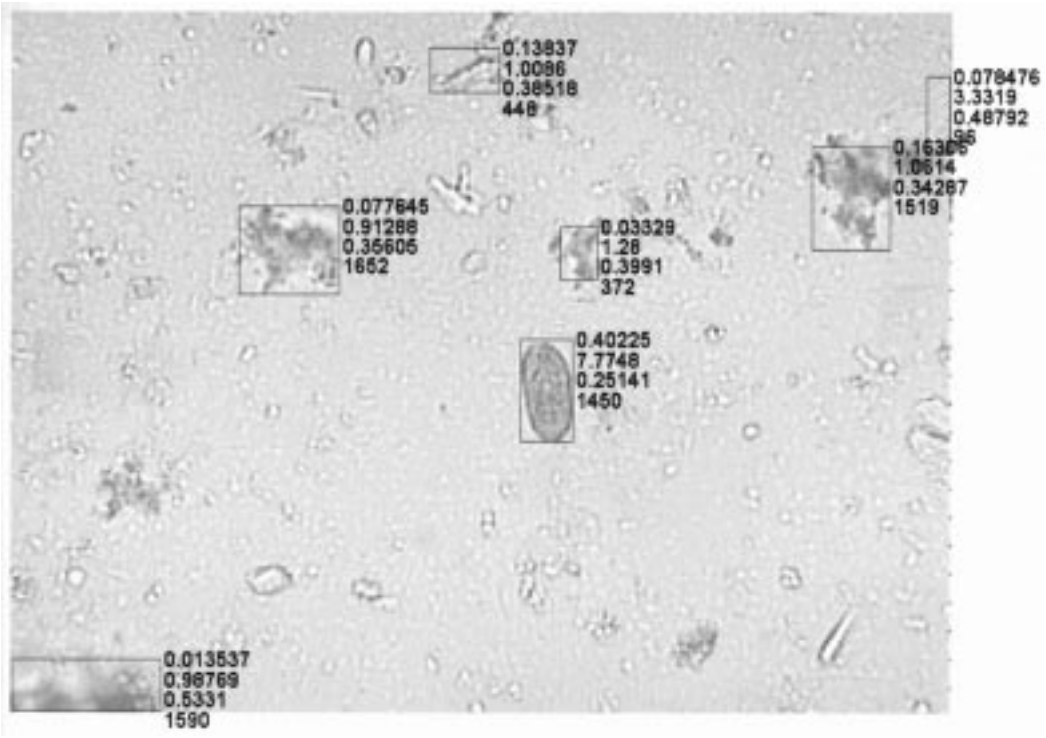
(a)

(b)

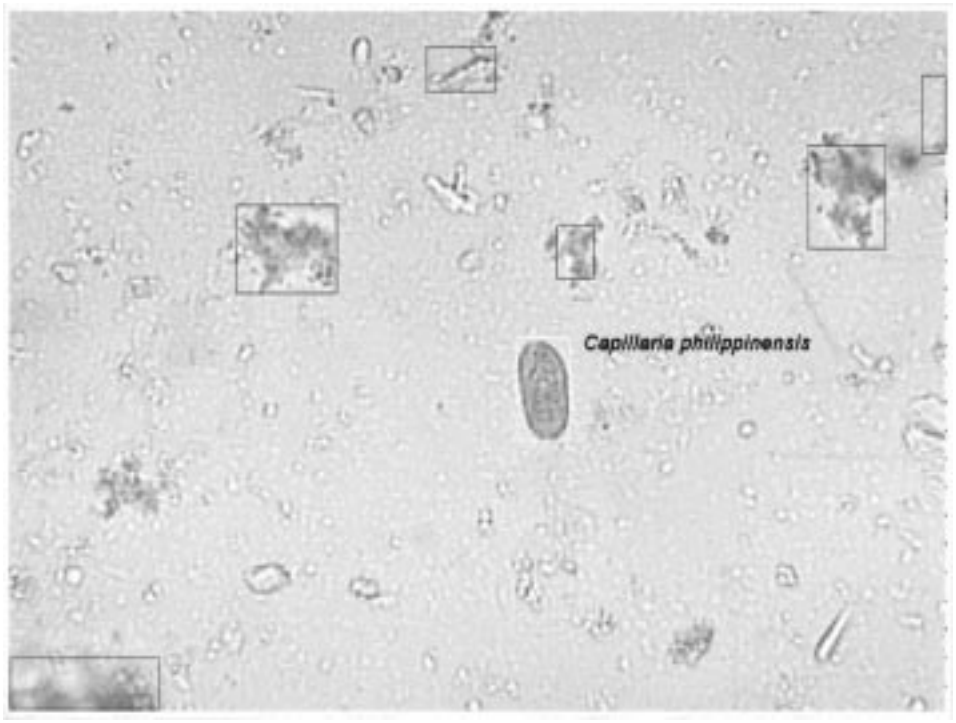
Fig. 9. Typical results of ANN-1 and ANN-2: (a) detection result of *C. sinensis*. (b) Identification result of (a).

a relationship between the optimal network topology and the available sample size in our specific application. The adequacy of the sample size of 82 still has not been verified. The proper size of data sample is highly related to the distributional characteristics of the given data samples. For example, in a case where the given data samples represent the classification problem well, relatively small number of data would provide satisfactory result.

In this study, ANN classification was performed by two sub classifiers. This separation made the training easier and reduced total processing time in real application to images. Most ANNs classify given patterns according to their geometric clustering in the pattern space [8]. In the case of fecal images, however, the artifacts are too dispersed in the feature space to be treated as one or few couples of classes. This would increase the complexity of ANN, that is, the



(c)



(d)

Fig. 9. (*Continued*.) Typical results of ANN-1 and ANN-2. (c) Detection result of *C. philippinensis* and (d) identification result of (c).

number of nodes and connections. Therefore, egg finding in the segmented objects was implemented separately from the identification of parasite species. The resultant two ANNs have much simpler architectures than they would have in one incorporated network. Moreover, this separation reduced the number of computations by skipping the application of ANN-2 to the artifacts preremoved by ANN-1.

The computed time for the overall procedure, including digital image processing and ANN classifiers was approximately 15 s/image. By discarding objects smaller than the minimum egg size at the segmentation, which eliminated unnecessary feature extraction, the computational load was alleviated. By the way, one specimen slide usually consists of 60 view fields under 200-fold magnified views. Therefore, if the developed algorithm is applied

to the entire section of slide plate, it will take about 15 min to process. This length of time, which appear rather lengthy, can be easily reduced by half or more by implementing routines optimally with professional programming languages and compilation.

The targets for the developed algorithm were seven typical human helminth eggs. For the identification of more parasitic species than in this study, a detailed analysis of materials inside the shell membrane [4] under 400-fold or higher magnification is indispensable. Much more features than employed in this study may be possible and even the best features, in terms of maximal variances and minimal redundancy, would be able to derived from them using dimensionality reduction methods like Karhunen–Loéve transformation [17]. Future work will accommodate this procedure.

The ultimate goal of this study is to develop an automatic examination system. Two different types of processing are possible. In an online processing system, the program performs the parasitic egg identification on one digital image before the next scan. In this case, total processing time should be reduced significantly. Using the large storage capacity of a modern PC, an off-line system scans the whole slide first and stores all digital images. Then, the program retrieves the stored images and carries out the identification. The developed algorithm will be an essential part of the main software of a completely automated system.

REFERENCES

[1] L. R. Ash and T. C. Orihel, *Atlas of Human Parasitology*, 3rd ed. Chicago, IL: ASCP, 1990.

[2] A. Joachim, N. Dulmer, and A. Daugschies, "Differentiation of two Oesophagostomum spp. from pigs, O. dentatum and O. quadrispinulatum, by computer-assisted image analysis of fourth-stage larvae," *Parasitol. Int.l*, vol. 48, pp. 63–71, 1999.

[3] A. Daugschies, S. Imarom, and W. Bollwahn, "Differentiation of porcine Eimeria spp. by morphologic algorithms," *Veterinary Parasitol.*, vol. 81, pp. 201–210, 1999.

[4] C. Sommer, "Quantitative characterization of texture used for identification of eggs of bovine parasitic nematodes," *J. Helminthol.*, vol. 72, no. 2, pp. 179–182, 1998.

[5] ——, "Digital image analysis and identification of eggs from bovine parasitic nematodes," *J. Helminthol.*, vol. 70, pp. 143–151, 1996.

[6] ——, "Quantitative characterization, classification and reconstruction of oocyst shapes of Eimeria species from cattle," *Parasitology*, vol. 116, pp. 21–28, 1998.

[7] R. C. Gonzalez, *Digital Image Processing*. New York: Addison-Wesley, 1993.

[8] J. A. Freeman and D. M. Skapura, *Neural Networks: Algorithms, Applications, and Programming Techniques*. New York: Addison-Wesley, 1992.

[9] R. J. Sohalkoff, *Pattern Recognition: Statistical, Structural and Neural Approaches*. New York: Wiley, 1992.

[10] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles.* New York: Addison-Wesley, 1981.

[11] K. Fukunaga, *Introduction to Statistical Pattern Recognition.* New York: Academic, 1990.

[12] R. N. Czerwinski, D. L. Jones, and W. D. O'Brien, "Line and boundary detection in Speckle images," *IEEE Trans. Image Processing*, vol. 7, pp. 1700–1714, Dec. 1998.

[13] C. W. Chen, J. Luo, and K. J. Parker, "Image segmentation via adaptive K-Mean clustering and knowledge-based morphological operations with biomedical applications," *IEEE Trans. Image Processing*, vol. 7, pp. 1673–1683, Dec. 1998.

[14] J. M. Gauch, "Image segmentation and analysis via multiscale gradient watershed hierarchies," *IEEE Trans. Image Processing*, vol. 8, pp. 69–79, Jan. 1999.

[15] M. A. Kupinski and M. L. Giger, "Automated seeded lesion segmentation on digital mammograms," *IEEE Trans. Med. Imag.*, vol. 17, pp. 510–517, Aug. 1998.

[16] N. Shareef, D. L. Wang, and R. Yagel, "Segmentation of medical images using LEGION," *IEEE Trans. Med. Imag.*, vol. 18, pp. 74–91, Jan. 1999.

[17] M. D. Srinath, P. K. Rajasekaran, and R. Viswanathan, *Introduction to Statistical Signal Processing with Applications.* Englewood Cliffs, NJ: Prentice-Hall, 1996.

Yoon Seok Yang (S'00) received the B.S. degree in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1996 and M.S. degree in interdisciplinary course (biomedical engineering major) from Yonsei University, Seoul, Korea, in 1998. He is currently a Ph.D. degree candidate in the Medical Electronics Laboratory (MELab) of the Seoul National University.

From 1996 to 1998, he worked as a Research Member at the Department of Biomedical Engineering of Yonsei University Medical Center. His main research topic was an early detection of breast cancer in digital mammogram. In 1998, he joined the Medical Electronics Laboratory (MELab) of the Seoul National University, where the major research activities are development of portable electronics and intelligent algorithms for medical and biological applications such as artificial pancreas, biosensors, telemedicine, and man-machine interface. His interests include signal processing. From 1999 to 2000, he worked on the ultrasonic biologic tissue characterization project funded by MEDISON, Korea, that develops various ultrasonic medical equipments.

Mr. Yang is a student member of Korea Society of Medical and Biological Engineering (KOSOMBE).

Duck Kun Park received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 2000. He is a M.S. degree candidate in the Medical Electronics Laboratory (MELab) of the Seoul National University.

Since 2000, he has been working in MELab. His interests include man-machine interface, augmented reality, neuroscience, bioMEMS, wireless portable health care system. He helped the production of EOG-based glasses type wireless man-machine interface. He has developed a game interface using EMG and acceleration together.

Mr. Park is a student member of Korea Society of Medical and Biological Engineering (KOSOMBE).

Hee Chan Kim (M'95) received the Ph.D. degree in control and instrumentation engineering (biomedical engineering major) from Seoul National University, Seoul, Korea, in 1989.

From 1982 to 1989, he has worked as a Research Member at the Department of Biomedical Engineering of Seoul National University Hospital. From 1989 to 1991, he was a Staff Engineer working on National Institute of Health (NIH) funded electro-hydraulic total artificial heart project at the Artificial Heart Research Laboratory of the University of Utah, Salt Lake City, USA. He joined the faculty of the Department of Biomedical Engineering, College of Medicine, Seoul National University in 1991, where he is currently an Associate Professor. He worked with the University of Utah Group again at the Department of Pharmaceutics and the Artificial Heart Research Laboratory as a Visiting Professor from 1993 to 1994. He is now leading the Medical Electronics Laboratory (MELab) of the Seoul National University, where the major research activities are development of intelligent algorithms and electronic instrumentations for medical and biological applications including artificial organs such as artificial heart and artificial pancreas, biosensors, telemedicine, and man-machine interface. In these areas, he has published over 38 peer-reviewed scientific papers in international journals.

Dr. Kim is a member of Korea Society of Medical and Biological Engineering (KOSOMBE), Institute of Electrical and Electronics Engineers/Engineering in Medicine and Biology Society (IEEE/EMBS), and American Society of Artificial Internal Organs (ASAIO).

Min-Ho Choi received the Ph.D. degree in parasite biochemistry (medical parasitology major) from Seoul National University, Seoul, Korea, in 1996.

From 1989 to 1993, he has worked as a Research Assistant at the Department of Parasitology, Seoul National University College of Medicine, Seoul Korea. He was a Public Health Doctor of National Institute of Health, Korea, from 1993 to 1996, where he counseled HIV-infected persons in Korea. He joined the faculty of the Department of Parasitology, Seoul National University College of Medicine in 1996, where he is currently an Assistant Professor. He is now a Visiting Scholar at the Department of Pathology, Division of Infectious Diseases, University of California, San Diego Medical Center, San Diego, CA, and he is studying the pathogenesis of amebiasis and transgenic amebae.

Dr. Choi is a member of Korea Society for Parasitology and the American Society of Parasitology.

Jong-Yil Chai received the M.D. and Ph.D. (medical parasitology) degrees from Seoul National University College of Medicine, Seoul, Korea, in 1976 and 1984, respectively.

He has been with the Department of Parasitology, Seoul National University College of Medicine since 1976. He is now Professor and Chairperson of the Department. For abroad studies, he worked on mucosal immunity to protozoan parasites, from 1986 to 1987, in Animal Parasitology Institute, US Department of Agriculture, Beltsville, MD, as a Visiting Scientist. He also studied on immunity to intestinal parasites, in 1993, in the Department of Parasitology, Institute of Medical Science, The University of Tokyo, Japan, as a Visiting Professor. He has published over 300 peer-reviewed academic papers in the field of medical parasitology and one textbook on clinical parasitology, and contributed many chapters of textbooks on clinical medicine. His research area includes a wide range of parasites of human beings including protozoa and helminths, but he has been mostly interested in the researches on intestinal flukes.

Dr. Chai is currently the President of Korean Society for Parasitology, and member of American Society of Tropical Medicine and Hygiene, American Society of Parasitologists, Society of Protozoologists, Society for Mucosal Immunology, and Helminthological Society of Washington, Royal Society of Tropical Medicine and Hygiene, and Japanese Society of Parasitology.