2014 22nd International Conference on Pattern Recognition

A Web-Based Education System for Reading Video Capsule Endoscopy

Hai Vu*, Yukiko Yanagawa¶‡, Tomio Echigo†, Masatsugu Shiba§, Hirotoshi Okazaki§, Yasuhiro Fujiwara§, Tetsuo Arakawa§, Yasushi Yagi‡,
*International Research Institute MICA Hanoi University of Science and Technology
†Dept. of Engineering Informatics, Osaka Electro-Communication University
‡The Institute of Scientific and Industrial Research, Osaka University
§Department of Gastroenterology, Osaka City University Graduate School of Medicine
¶Technology Development Center, OMRON Corporation

***Abstract*—The interpretive skills of medical doctors and medical technologists who examine video capsule endoscopy (VCE) in clinical practice are usually improved through hands-on courses. Such courses require that a large volume of cases be undertaken as part of the training, and they thus consume a considerable amount of the trainees' time. This paper describes an e-learning system that reduces the training time in addition to enhancing the quality of the educational process with regard to reading VCE. To achieve this goal, we focused on organizing training courses in order to appropriate for the laborious conditions that exist when reading VCE. The designed courses help the trainees acquire knowledge of abnormal regions and become familiar with reading VCE before taking examinations under conditions similar to those actual clinical practice. The proposed training modality was developed as an e-learning application on the World Wide Web. Thus, it can be easily extended to a wide range of trainees. In the experiments, 20 participants completed the self-learning training procedures in approximately 3 hours. The proposed system is much faster than conventional hands-on courses, which require a minimum of 8 hours. Furthermore, the trainees' learning performances in the final examinations confirmed that the proposed system is particularly effective for inexperienced examining doctors.**

I. INTRODUCTION

Achieving competency in reading video capsule endoscopy (VCE) [1] is a crucial task for examining doctors (medical doctors or medical technologists) who examine VCE sequences in clinical practice. Although VCE technology has emerged as an important tool for specialists in finding suspected lesions in the human small bowel, such as obscure gastrointestinal (GI) bleeding and Crohn's disease [2], [3], [4]. However, though this new imaging modality represents a breakthrough technique, examining doctors often have to work under laborious conditions when reading the VCE. They have to exert great concentration as they check consecutive frames. Many mucosal lesions appear as mere artifacts among the intestinal structures and are displayed relatively fast. They may be incomplete in appearance or may even be invisible because only a subregion is captured [5]. It appears clear that examining doctors need to improve their reading skills through training programs. For example, an investigation by Sidhu et al. in [6] among capsule endoscopists in the United Kingdom showed that prior medical knowledge enabled experienced doctors to interpret VCE more accurately than novices, such as medical students and medical technicians lacking the appropriate experience. The additional requirements of training programs have been discussed in various medical reports [7], [8], [6], [9].

Most conventional training approaches are based on hands-on or on-the-job training courses. The interpretive skills of trainees are increased under the guidance of experienced colleagues in the daily clinic. Guidelines of the American Society for Gastrointestinal Endoscopy (ASGE) [10] recommend that a trainee should complete a hands-on course with a minimum of 8 hours' continuous training. Shaver et al. [11] investigated the learning curves of trainees in reading VCE. They found that junior and intern endoscopists were able to gain moderate accuracy in detecting lesions after examining the sequences from 20 cases. Other studies, such as that of Lo [5], indicated that supervised capsule endoscopy with a minimum of 25 to 50 cases is required. With such training, there is a lack of structured materials and evaluation functions, and so it is important that not only novices but also experienced gastroenterologists should read as many cases as possible. Obviously, hands-on training courses demand a great deal of the trainees' time. Recently, Erber [9] recommended that the deficiency of formal training programs has to be addressed to enhance the accuracy in reading VCE. Guidelines need to be established that build upon the trainees' medical knowledge; further tasks involve formal certification of the skills of examining doctors with regard to the detection and categorization of relevant diseases. Postgate et al. [12] made a notable evaluation of the efficiency of a computer-based training program for reading VCE. Their method follows a scheme of pre-training and post-training performance evaluation.

In light of the above, we propose in this paper a training modality for reading VCE. The intention with the proposed modality is to reduce the training time for trainees, particularly among inexperienced examining doctors. Toward this end, we will focus on organizing a structured curriculum for the learning stages. Our learning algorithms start with training courses that have very short image sequences so that trainees are able to easily learn and acquire knowledge of the characteristics of various abnormalities. To meet the demands of actual VCE reading, the training procedures continues with

1051-4651/14 $31.00 © 2014 IEEE
DOI 10.1109/ICPR.2014.470

a series of longer image sequences. The main purpose of this stage is that trainees become familiar with the changing images in VCE, such as in intestinal contractions, where there are sudden changes in the capsule direction. After each learning stage, trainees' performances are evaluated to confirm that they understand what they have seen and are able to describe their findings using standardized terminology. The learning algorithms therefore contain loops of trials until the trainees' performances surpass the baseline levels. Consequently, trainees are able to reliably identify and interpret different pathologies at the final examination stage, which is similar to the actual conditions in reading VCE. The main advantages of the proposed method are that trainees are able to pass the examinations after approximately 3 hours under self-learning, which is much faster than the 8 hours' minimum requirements of hands-on training courses.

There are two main differences between our own approach and that developed by Postgate et al. [12]. First, beyond the evaluation study of Postgate et al., in this paper we provide details of a proper training system for reading VCE. With the proposed system, we aim to reduce the training time as well as enhance the quality of the training with respect to reading VCE. Therefore, the content of our training procedure is different from that of the training module proposed by Postgate et al. [12], whose purpose was only that of evaluation. Second, the organization of the training procedures in our proposed system is clearly adapted to the actual conditions of VCE reading. This was indicated as a weak point in the study of Postgate et al. by the authors themselves [12] (see discussion in [12]). Furthermore, our proposed system can be extended to a wide range of users. This represents a further development from the stand-alone application described by Postgate et al. [12] to a Web-based application. Our system is the first e-learning system to be reported for VCE technology. The rest of the paper is structured as follows: Section II describes the proposed training modality. Setting up the proposed system and experimental results are reported in section III. Finally, section IV concludes key points, and suggest directions for future studies.

II. The proposed training modality

The proposed learning algorithms consist of a series of training courses. These courses are made up of three training stages: (1) gaining knowledge of the abnormal regions; (2) becoming familiar with conditions relating to reading VCE; and (3) verifying medical knowledge. At each stage, there is a series of trials that cover the full scale of abnormal regions. A training stage is repeated until the trainee's performance exceeds baseline levels. To complete the training courses, trainees have to pass all stages from (1) to (3). Before explaining the proposed learning algorithms in detail, we will introduce features of the learning system that support trainees throughout the entire training.

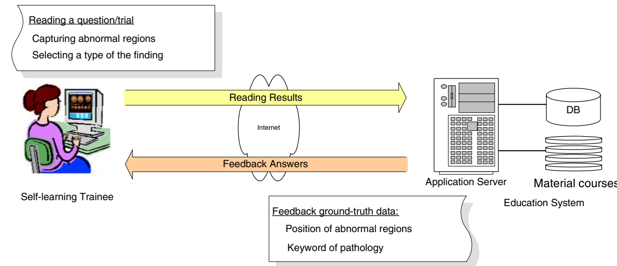

Fig. 1. One person using the proposed system

A. Learning styles of the proposed system

The proposed modality is an e-learning application that supports self-learning or constitutes a one-person learning scheme. Trainees are required to follow the content without direct advice or guidance from experienced colleagues. The learning activities are shown in Fig. 1. Trainees read or scan an image sequence and then detect the positions of the abnormal region. A multiple-choice question with one correct answer among four possible options is given for each finding position. Before moving to the next question, the system automatically feeds back ground-truth results so that trainees are able to verify their answers (see the screenshot in Fig. 5). This scheme provides trainees try-and-error training, which is similar to an on-the-job training scheme without the guidance of experts. In terms of repetitive learning, each kind of abnormal region is repeated in several sequences in the training stages. Furthermore, the proposed system accommodates incorrect answers by the trainee. The incorrect cases are repeatedly displayed until the trainee is able to find the abnormal positions as well as provide a correct answer for the type of abnormal region.

B. Organization of the training courses

The course material contains a series of trials (video sequences), the preparation of which initially demands much greater consideration and is more time-consuming than with most hand-on courses, and the trials also require the input of very experienced practitioners. The trials are organized in three training stages: *Learning, Training*, and *Examination*, as shown in Fig. 2. First, the trainee needs to acquire medical knowledge, such as that relating to the characteristics of abnormal regions, by reading short trials. This stage-*Learning*- is similar to that with atlas-based learning approaches, in which each type of abnormal region is constructed repetitively in several short sequences. There are 50 frames per sequence. In the *Training*, the course material contains a set of long sequences. There are three repetitive procedures: *Training #1, Training #2, and Training #3*. The length of the sequences in this procedure is 200 frames. With the use of longer sequences, trainees are able to enhance their skills in finding abnormal regions under changing image conditions, such as capsule movements and intestinal contractions appearing in the image sequences. The position of a certain abnormal region in a sequence of the training trials is related to the corresponding

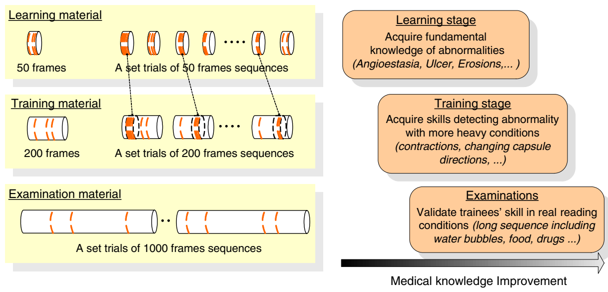

Fig. 2. Organization of the course material. The purpose of each corresponding stage is indicated

sequence in the learning trials, as shown in Fig. 2. This ensures that a trainee can more easily recognize an abnormal region that has already been recognized in the *Learning* stage.

Finally, the *Examination* contains sequences that consist of 1000 frames (or approximately 8 minutes' transiting of the capsule device in the GI tract). The image changes in these trials are close to those of normal conditions when reading VCE. Several different types of abnormal regions can appear in an examination sequence. No new types of abnormal regions are introduced in this procedure. However, unlike in the *Training*, the material in this stage consists of new sequences, which did not display in the previous trials.

C. Learning performance assessments

In the proposed system, trainees are required to understand what they see and describe their findings using standardized terminology. Therefore, two criteria are used to evaluate the learning performance:

- Sensitivity α of the captured abnormality: This parameter measures how a trainee can detect abnormal regions in trials in a training stage.. This parameter is defined by:

$$\alpha = \frac{TP}{TP + FN} 100(\%) \tag{1}$$

  TP: number of true positive finding; FN: number of false negative finding
- Accuracy β of the captured frames: Because interpreting VCE requires understanding the type of abnormal regions, the β measures how a trainee chooses the correct answers from the multi-choice questions. This parameter is defined by:

$$\beta = \frac{\text{Number of corrected answers}}{\text{Total of the trials in a training stage}} 100(\%) \tag{2}$$

As part of a formal training program, learning objectives are evaluated through baseline levels that indicate threshold numbers to assure competency in reading VCE. In the proposed system, the baseline levels of two parameters (α,β) are identified in advance. Three medical doctors (MDs) with extensive experience in reading VCE provide their evaluations of the course material. Their readings are used for the baseline levels. These values are given in Table I.

TABLE I
THE BASELINE LEVELS IN EACH TRAINING PROCEDURE AS IDENTIFIED BY EXPERIENCED MDS

Criteria	*Learning*		*Training*		*Examination*	
	α	β	α	β	α	β
Experienced MDs	80.5 %	88.7 %	65.0 %	88.1 %	44.9 %	86.0 %
Baseline	80.0 %	80.0 %	70.0 %	80.0 %	50.0 %	80.0 %

D. Learning procedures

The learning procedure algorithm is shown in Fig. 3. A trainee starts the *Learning* by checking abnormal regions in short sequences. The performance of the trainee is evaluated by sensitivity α and accuracy β. If these parameters are above the baseline levels, which are given in Table I, e.g., $\alpha \geqslant 80\%$ and $\beta \geqslant 80\%$, the trainee advances to the next course in the *Training*. There are three repetitive courses at this stage. The learning performances necessary to pass such stage are $\alpha \geqslant 80\%$ and $\beta \geqslant 70\%$. Finally, in the *Examination*, trainees receive authorization when their performances are $\alpha \geqslant 80\%$ and $\beta \geqslant 50\%$. A stage is repeated if a trainee's performance is below the corresponding baseline level.

III. SYSTEM DEVELOPMENT AND LEARNING RESULTS

A. System development

The proposed training modality is developed in the form of an e-learning application through the World Wide Web. The main components of the proposed system are shown in Fig. 4. On the server side, the application server including classes is developed using Java Servlet technology [13]. A Web server (Tomcat Server Ver. 6.0 [14]) provides an application programming interface (API) to work with these classes. The Java Servlet classes implement the dedicated functions of the proposed learning algorithms, such as control learning stages, database management, and response requests from clients to verify answers. A database (using MySQL Ver. 5.0) handles the organization of materials, trainee information, and logging trainee activities. Videos of the course material are also stored on a file server. The client side is made up of graphical user interfaces (GUIs) created using an XML-based user interface markup language (MXML). User event handles are written in Action Scripts language. The GUIs support the main functions of the e-learning applications, such as image sequence scanning, capturing abnormal regions, and showing answers to trainees. A screenshot of the GUI of the proposed e-learning system is shown in Fig. 5

B. Learning results

Our e-learning application uses material collected from a database of 300 patients at Osaka City University Hospital. The database covers a wide range of age, gender, and type of abnormal finding. Twenty participants, including 17 trainees and three experienced MDs, participate in the training courses of the proposed e-learning system. Each participant is assigned an account (including user ID and password) and is asked

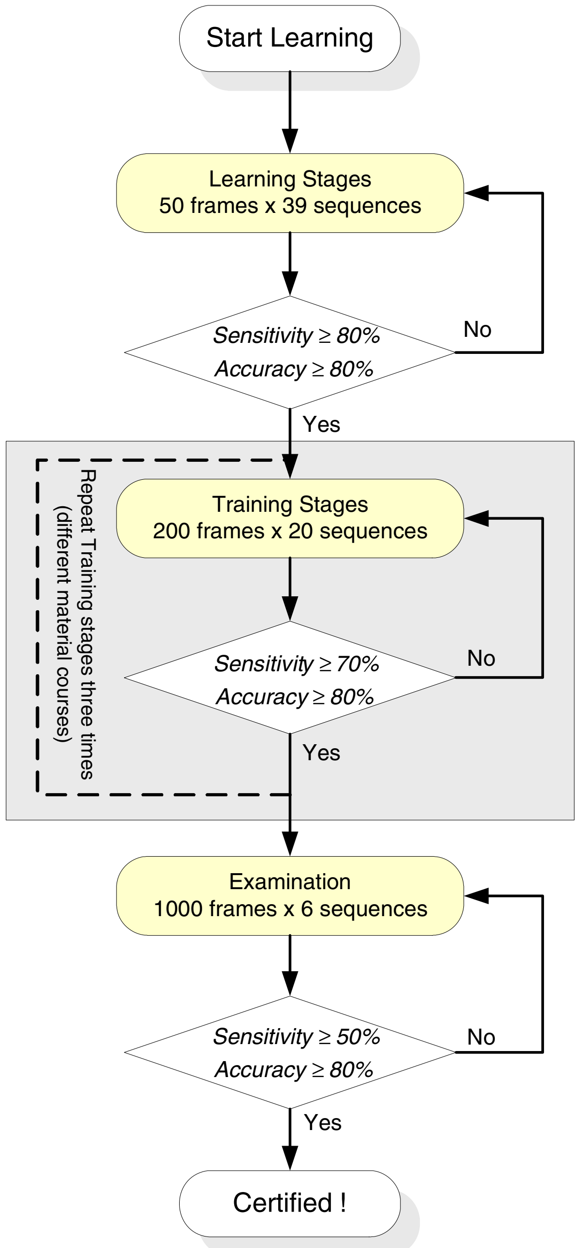

Fig. 3. The proposed learning algorithms

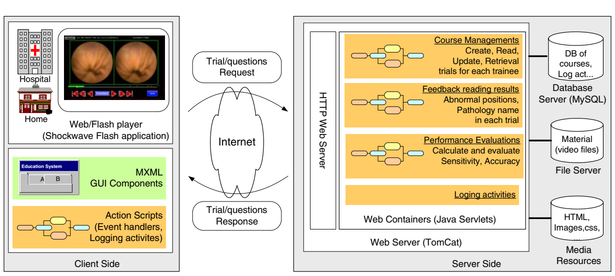

Fig. 4. The main components of the e-learning system

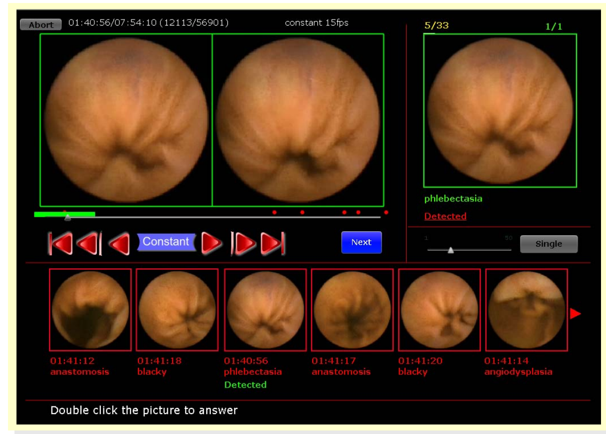

Fig. 5. A screenshot of the GUI of the proposed system

to complete the training programs under self-learning. The learning algorithms utilize the baselines identified in Table I. To quantity the learning performances of the trainees, we measure the sensitivity (α) and accuracy (β). The evaluations compare the trainees' performances with that of the experienced MDs. The main purpose of the evaluation is to confirm that the trainees have adequate medical knowledge to detect and categorize abnormal regions under the actual conditions of reading VCE.

Table II reports the performance of trainees measured in terms of sensitivity (α) and accuracy (β). They are the results of four training procedures before the final examinations: *Learning*, *Training #1*, *Training #2*, and *Training #3*. With each result, the trainees' performance in the first trial and last trial are reported (in the form of the ratio *a/b*). As can be seen, only a few trainees (such as ID #7 and ID #8) were able to attain the baseline level the first time they attended a course. Most trainees had to take a training course several times before they were able to pass. The data in Table II clearly confirm that all 17 trainees had successfully completed the training course such that they could take the final examinations. We compare the trainees' performances from the last trials (b values in Table II) with the results of the experienced MDs. Statistical measurements: average and range values $[min, max]$ of sensitivity and accuracy are shown in Fig. 6 and Fig. 7, respectively. Based on these comparisons, after the completion of the training courses, the interpretive skills of the trainees when reading VCE were approximately at the level of the experienced MDs. In other words, the trainees were able to achieve competency in reading VCE under self-learning.

We then compared the training time of the proposed method with the minimum requirements of hands-on training courses. As noted in the review by Erber. [9], hands-on training courses require a minimum of 8 hours. The training time of trainees using the proposed system is given in Table III. As indicated, the total training time was 3 hours on average 194 $\pm$ 56 min. This time is much less than with conventional hands-on

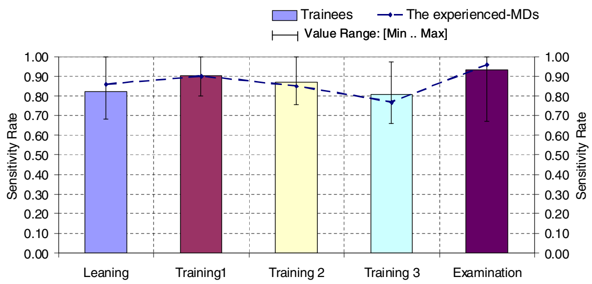

Fig. 6. Comparison of sensitivity (α) between the trainees and experienced MDs in each training procedure

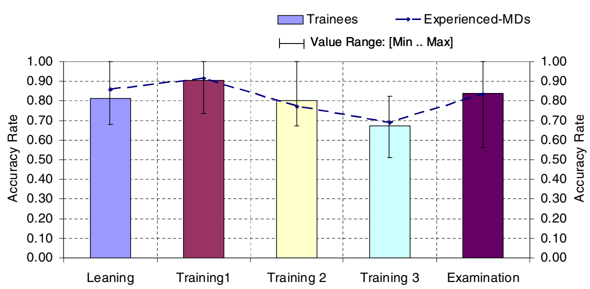

Fig. 7. Comparison of accuracy (β) between the trainees and experienced MDs in each training procedure

courses. These results confirm the effectiveness in reducing the training time with the proposed e-learning system. The proposed system was deployed in an internal environment with support from medical doctors and medical technologists in Osaka City University Hospital, Japan., We are currently uploading the proposed system to a public Web site. Interested individuals can register as members and join the training program at this Web site: www.ce-elearning.org.

As shown in Table III, experienced MDs needed an average of only 140 ± 27 min. to finish the training procedures. This was much less than the training time of the 17 trainees. The group of 17 trainees required an average of 70 $\pm$ 22 min. to complete the *Learning*, whereas experienced MDs needed only 30 $\pm$ 16 min.; such differences can also be observed in *Training #1* and *Training #2*. The main reason for this difference is that trainees lack experience with abnormalities at the *Learning* stage. Their medical knowledge shows improvement after *Training #1, #2, and #3*: evidence for this is seen in the training time for *Training #3* and in the *Examinations* results. Almost all trainees were able to pass the final examinations, and their results were the same as those of the experienced MDs. This indicates that the proposed system is effective for such trainees as inexperienced MDs.

IV. Conclusions

This paper presents the first proper training program for reading VCE. The most notable feature of the proposed system is that it significantly reduced the training time of trainees. Twenty participants took 3 hours to pass the final examination in the experiments. This training time is much less than the 8 hours required in conventional hands-on training courses. To achieve this, the proposed system takes into account the organization of courses and evaluates learning performance functions. Through a series of training courses, trainees acquire knowledge of abnormal regions and become familiar with the actual conditions in reading VCE. The proposed system was developed as an Web-based learning application. In the future, larger trainee groups will be used with the proposed training system for further evaluation.

V. Acknowledgements

The first author thanks to Vietnam National Foundation for Science and Technology Development (NAFOSTED) for their supports to fulfill the study.

References

[1] G. Iddan, G. Meron, A. Glukovsky, and P. Swain, "Wireless capsule endoscope," *Nature*, vol. 405, p. 417, 2000.

[2] P. Swain and A. Fritscher-Ravens, "Role of video endoscopy in managing small bowel disease," *GUT*, vol. 53, pp. 1866–1875, 2004.

[3] B. Lewis, R. Franchis, and G. Gay, "Consensus report," in *Proc. of International Conference on Capsule Endoscopy*, 2006.

[4] G. I. Homepage, "http://www.givenimaging.com /cultures/en-us/given/english," Oct 2007.

[5] S. Lo, "How should we do capsule reading?" *Techniques in GastroIntestinal Endoscopy*, vol. 8, no. 4, pp. 146–148, 2006.

[6] R. Sidhu, P. Sakellariou, M. McAlindon, J. Leeds, and et al., "Is formal training necessary for capsule endoscopy? the largest gastroenterology trainee study with controls," *Digestive and Liver Disease*, vol. 40, pp. 298–302, 2008.

[7] Y. Niv and G. Niv, "Capsule endoscopy examination - preliminary review by a nurse," *Digestive Diseases and Sciences*, vol. 50, no. 11, pp. 2121–2124, 2005.

[8] N. Bamji, B. Bosworth, R. Sockolow, and et al., "Lack of education and training in video capsule endoscopy among pediatric gastroenterology fellows," *American Journal Gastroenterol*, vol. 102, no. 9, p. AB 1106, 2007.

[9] J. Erber, "Wireless capsule endoscopy: where and how to learn?" *GastroIntestinal Endoscopy*, vol. 68, no. 1, pp. 115–117, 2008.

[10] A. S. for Gastrointestinal Endoscopy, "Endoscopy by nonphysicians," *GastroIntestinal Endoscopy*, vol. 69, no. 4, pp. 767–770, 2009.

[11] C. Shaver, J. Rivera, J. McKinley, and P. Brady, "Capsule endoscopy learning curve," *GastroIntestinal Endoscopy*, vol. 59, no. 5, p. 123, 2004.

[12] A. Postgate, A. Haycock, S. Thomas-Gibson, A. Fitzpatrick, P. Bassett, S. Preston, P. B. Saunders, and C. Fraser, "Computer-aided learning in capsule endoscopy leads to improvement in lesion recognition ability," *GastroIntestinal Endoscopy*, vol. 70, no. 2, pp. 311–316, 2009.

[13] J. S. T. Overview, "http://www.oracle.com/technetwork/java/javaee/servlet/index.html/," Nov 2011.

[14] T. S. Hompage, "http://tomcat.apache.org/," Nov 2011.

TABLE II
SENSITIVITY α AND ACCURACY β MEASUREMENTS * OF THE PARTICIPANTS.

ID	*Learning*		*Training #1*		*Training #2*		*Training #3*		*Examination*	
	α(%)	β(%)	α(%)	β(%)	α(%)	β(%)	α(%)	β(%)	α(%)	β(%)
The trainee results										
1	82/92	89/94	93	92	88	94	59	86	89	91
2	74/95	75/94	95	93	56/88	63/93	29/77	60/88	67	80
3	82	98	100	100	100	100	78	90	85	100
4	89	91	88	86	78	88	65	86	80	89
5	76/92	79/94	88	85	61/89	69/97	81	97	73	85
6	71/86	88/95	85/95	85/93	67	77	57/88	87/89	78	100
7	97	97	100	100	93	100	82	93	88	100
8	96	100	95	95	100	100	57/82	76/97	91	100
9	61/94	79/95	82	81	75	88	61/78	91/90	62	93
10	63/97	89/97	89	88	83	93	68/93	83/97	100	100
11	50/87	54/89	73	80	43/95	56/95	71	89	75	89
12	71/92	88/92	100	100	77	93	50/71	56/88	100	100
13	68/89	76/89	90	89	61/100	64/100	32/74	55/85	50/63	49/85
14	61/84	66/91	77	80	56/82	65/93	43/74	52/80	88	86
15	45/92	50/97	100	100	82	82	46/77	68/83	100	100
16	63/87	68/90	88	87	71	83	41/70	52/84	100	100
17	74	85	93	90	73	85	43/89	63/85	63	88
The experienced-MDs Results										
18	74/92	74/95	87	86	72	81	71	74	88	100
19	100	100	94	95	85	89	70	82	100	100
20	63/89	67/92	90	91	85	86	43/85	43/92	88	92

* *a/b*: In each training procedure, the results at the first trial *(a)* and last trial *(b)* are reported. If only *a* is reported, this means that the trainee needed only one trial to attain (or surpass) baseline levels.

TABLE III
THE TRAINING TIME (IN MIN) OF THE PARTICIPANTS

Trainee ID.	*Learning* (min)	*Training #1* (min)	*Training #2* (min)	*Training #3* (min)	*Examination* (min)	Total (min)
The trainee results						
1	33	8	14	14	23	92
2	104	22	31	79	23	258
3	65	69	41	66	18	260
4	43	10	19	21	26	119
5	110	18	50	71	44	294
6	80	19	11	46	17	175
7	42	20	24	34	18	138
8	76	19	10	37	21	162
9	85	10	17	71	22	205
10	64	16	16	46	13	153
11	82	21	47	14	22	186
12	70	18	13	77	32	210
13	72	20	40	56	56	271
14	65	21	39	58	32	215
15	69	11	27	64	21	192
16	35	18	15	69	12	149
17	92	17	33	41	41	224
Avg.$\pm$Std.	70 ± 22	20 ± 13	26 ± 13	51 ± 21	26 ± 12	194 ± 56
The experienced-MDs results						
18	19	28	11	41	34	133
19	22	13	15	44	23	117
20	48	13	15	76	17	169
Avg.$\pm$Std.	30 ± 16	18 ± 9	14 ± 2	54 ± 19	25 ± 8	140 ± 27