

ICFHR 2016 Competition on the Analysis of Handwritten Text in Images of Balinese Palm Leaf Manuscripts

Jean-Christophe Burie, Mickaël Coustaty, Setiawan Hadi, Made Windu Antara Kesiman, Jean-Marc Ogier, Erick Paulus, Kimheng Sok, I Made Gede Sunarya, Dona Valy

▶ **To cite this version:**

Jean-Christophe Burie, Mickaël Coustaty, Setiawan Hadi, Made Windu Antara Kesiman, Jean-Marc Ogier, et al.. ICFHR 2016 Competition on the Analysis of Handwritten Text in Images of Balinese Palm Leaf Manuscripts. 15th International Conference on Frontiers in Handwriting Recognition 2016, Oct 2016, Shenzhen, China. pp.596-601, 10.1109/ICFHR.2016.107 . hal-01389856

HAL Id: hal-01389856
https://hal.science/hal-01389856

Submitted on 30 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ICFHR 2016 Competition on the Analysis of Handwritten Text in Images of Balinese Palm Leaf Manuscripts

Jean-Christophe Burie[1], Mickael Coustaty[1], Setiawan Hadi[3], Made Windu Antara Kesiman[1,2], Jean-Marc Ogier[1], Erick Paulus[3] , Kimheng Sok[4], I Made Gede Sunarya[2] and Dona Valy[4],

[1]Laboratoire Informatique Image Interaction (L3i), University of La Rochelle
Avenue Michel Crépeau 17042, La Rochelle Cedex 1, France
{jean-christophe.burie, mickael.coustaty, made_windu_antara.kesiman, jean-marc.ogier}@univ-lr.fr
[2]Laboratory of Cultural Informatics (LCI), Ganesha University of Education
Jalan Udayana No 11, 81116, Singaraja, Bali, Indonesia
[3]Robotics, Artificial Intelligence, and Digital image (RAID) Laboratory,
University of Padjajaran, Bandung, Indonesia
[4]Department of Information and Communication Engineering,
Institute of Technology of Cambodia, Cambodia

Abstract — **This paper presents the results of the Competition on the Analysis of Handwritten Text in Images of Balinese Palm Leaf Manuscripts that was organized in the context of the 15th International Conference on Frontiers in Handwriting Recognition (ICFHR-2016). This competition provides a suitable challenge for testing and evaluation of robustness for some methods, image features and descriptors which were already proposed for handwritten text analysis of document image. In this competition, three different challenges in document analysis of palm leaf manuscript images are proposed: Challenge 1: Binarization of Palm Leaf Manuscript Images, Challenge 2: Query-by-Example Word Spotting on Palm Leaf Manuscript Images, and Challenge 3: Isolated Character Recognition of Balinese Script in Palm Leaf Manuscript Images. The first handwritten Balinese palm leaf manuscript dataset, the *AMADI_LontarSet*, is used for performance evaluation. This paper describes the competition details including the dataset creation and the ground truth construction, the evaluation measures used, a short description of each participant as well as the performance of the all submitted methods.**

Keywords - Balinese script; palm leaf manuscript; binarization, query-by-example word spotting, isolated character recognition, performance evaluation

I. INTRODUCTION

Palm leaf manuscripts offer a new challenge in document image analysis system due to the physical characteristics and conditions of the manuscripts. The palm leaf manuscripts contain discolored parts and artefacts due to aging and low intensity variations or poor contrast, random noises, and fading. Several deformations in the character shapes are visible due to the merges and fractures of the use of nonstandard fonts, varying space between letters, and varying space between lines [1,2].

This competition tries to bring added value to digitized palm leaf manuscripts by developing methods to analyze and access quickly and efficiently to the content of manuscript and to make palm leaf manuscript more accessible, readable and understandable to a wider audience. In this competition, three different challenges in document analysis of palm leaf manuscript images are proposed: Challenge 1: Binarization of Palm Leaf Manuscript Images, Challenge 2: Query-by-Example Word Spotting on Palm Leaf Manuscript Images, and Challenge 3: Isolated Character Recognition of Balinese Script in Palm Leaf Manuscript Images.

The three challenges on this competition represent three problems which are still challenging and are still an open question on handwritten text analysis of document image especially for palm leaf manuscript images. The Balinese palm leaf manuscripts were written in Balinese script in Balinese language. In the scope of this competition, we will focus on scripts written on palm leaf whose characters are even more elaborate in ancient written form. The manuscript datasets used in this competition, the *AMADI_LontarSet*, are the sample images of the palm leaf manuscripts from Bali, Indonesia. In order to obtain the variety of the manuscript images, the sample images were collected from 23 different collections (contents), which come from five different locations (regions): two museums and three private families. It consists of randomly selected ten collections from Museum Gedong Kertya, City of Singaraja, Regency of Buleleng, North Bali, Indonesia, four collections from Museum Bali, City of Denpasar, South Bali, seven collections from a private family collection from Village of Jagaraga, Regency of Buleleng, and two others collections from a private family collections from Village of Susut, Regency of Bangli and from Village of Rendang, Regency of Karangasem. From those 23 collections, we captured 393 pages of palm leaf manuscript. The description of dataset for each challenge is presented in each section of the challenge. This competition provides a suitable challenge for testing and evaluation of robustness for some methods, image features and descriptors which were already proposed for handwritten text analysis of document image. All researchers were invited to participate in one or more challenges.

This paper is organized as follow: Section II gives a brief description about the participants. Section III-V present the detail description of each challenge. Finally, some conclusions are given in Section VI.

II. THE PARTICIPANTS

Eight research groups registered to this competition: seven groups registered to Challenge 1, three groups registered to Challenge 2, and six groups registered for Challenge 3. But only five research groups accomplished the task and submitted their results: four groups accomplished Challenge 1, and three groups accomplished Challenge 3. There is one group who submitted two methods for Challenge 3. Unfortunately, there is no group who accomplished Challenge 2. Brief introduction of the five research groups who submitted their results are as follows:

Group 1 (Challenges 1 with two methods and 3): Submitted by Yunxue Shao, Jiantao Zhou and Guanglai Gao, from College of Computer Science, Inner Mongolia University.

Group 2 (Challenge 1): Submitted by Chris Tensmeyer, from Computer Science Dept., Brigham Young University.

Group 3 (Challenges 1 and 3): Submitted by Su Bolan and Lu Shijian, from Institution of Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore and Zhang Xi, from Advanced Remanufacturing and Technology Centre (ARTC), Agency for Science, Technology and Research (A*STAR), Singapore.

Group 4 (Challenge 1): Submitted by Dr. Deepak Kumar, from Department of Electronics & Communication Engineering, Dayananda Sagar Academy of Technology and Management (DSATM), Bengaluru, Karnataka, India.

Group 5 (Challenge 3): Submitted by Cristinel Codrescu, from University of Salzburg.

Brief description of their methods are given in each section of the challenge.

III. CHALLENGE 1: BINARIZATION OF PALM LEAF MANUSCRIPT IMAGES

A. Description of the Challenge

This challenge is focused on the binarization method for palm leaf manuscript images. The binarization problem is still challenging and is still an open question especially for palm leaf manuscript images. Visually, several alternative well-known binarization algorithms do not provide a good binarized image for palm leaf manuscript images. They extract unrecognizable characters on palm leaf manuscripts with noise. Therefore, to separate text (and, if any, graphic) from the background of palm leaf manuscript images, a specific and adapted binarization technique is required. The input of this challenge is a set of original color images of palm leaf manuscript, and the output is the binarized image of the manuscript.

B. Dataset

To create the ground truth binarized image of palm leaf manuscript, we adopted a semi-automatics framework for construction of ground truth binarized image [2]. For our manuscript collection, the skeletonization process is completely performed by human. The skeletons of the Balinese character (and, if any, graphic) in palm leaf manuscript image were traced and drawn manually with PixLabeler tool [3]. The skeletons are then automatically processed to create the estimated ground truth binarized image of the manuscript. Human subjectivity has a great effect in producing a variability of ground truth binarized image. This phenomena becomes much more visible when we are working on ancient type manuscript which is still hard to be ground truthed even by human. Therefore, for our dataset, one manuscript was ground truthed by two different ground truthers. And for the binarization challenge in this competition, we will evaluate each binarized image with two different ground truth binary images. The dataset is partitioned into training and test subsets. For the training subset, we have provided 50 original images, 50 ground truth binary images Version 1 (from the 1st ground truther) and 50 ground truth binary images Version 2 (from the 2nd ground truther) (Fig.1). For the testing subset, 50 original images (different from the training subset) have been provided.

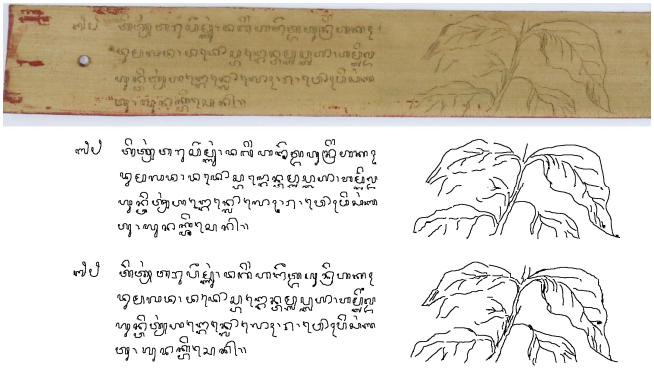


Figure 1. From top to bottom: original image, ground truth binary image version 1, ground truth binary image version 2

C. Performance Evaluation

Following our work in [1], three metrics of binarization evaluation which are used in the DIBCO 2009 contest [4], are used in the evaluation for this challenge. Those three metrics are F-Measure (FM), Peak SNR (PSNR), and Negative Rate Metric (NRM).

F-Measure (FM): FM is defined from Recall and Precision.

$$\mathrm{Re}\,call = \frac{TP}{FN+TP} * 100 \quad (1)$$

$$\mathrm{Pr}\,ecision = \frac{TP}{FP+TP} * 100 \quad (2)$$

TP, defined as true positive, occurs when the image pixel is labeled as foreground and the ground truth is also. FP, defined as false positive, occurs when the image pixel is labeled as foreground but the ground truth is labeled as background. FN, defined as false negative, occurs when the image pixel is labeled as background but the ground truth is labeled as foreground.

$$FM = \frac{2 * \mathrm{Re}call * \mathrm{Pr}ecision}{\mathrm{Re}call + \mathrm{Pr}ecision} \quad (3)$$

A higher F-measure indicates a better match.

Peak SNR (PSNR): PSNR is calculated from Mean Square Error (MSE).

$$MSE = \sum_{x=1}^{M} \sum_{y=1}^{N} \frac{(I_1(x,y) - I_2(x,y))^2}{M*N} \quad (4)$$

$$PSNR = 10 * \log_{10}(\frac{C^2}{MSE}) \quad (5)$$

where C is defined as 1, the difference between foreground and background colors in the case of binary image. A higher PSNR indicates a better match.

Negative Rate Metric (NRM): NRM is defined from negative rate of false negative (NR_{FN}) and negative rate of false positive (NR_{FP}).

$$NR_{FN} = \frac{FN}{FN+TP} \quad (6)$$

$$NR_{FP} = \frac{FP}{FP+TN} \quad (7)$$

TN, defined as true negative, occurs when both the image pixel and ground truth are labeled as background. The definitions of TP, FN, and FP are the same than the ones given for the F-Measure.

$$NRM = \frac{NR_{FN} + NR_{FP}}{2} \quad (8)$$

A lower NRM indicates a better match.

For each binarization metric evaluated on each ground truth binary images, we rank the participantôs score. We then assign point, from 1 (best result) to N (worst result, N is the number of participants). All points from 3 binarization metrics evaluated on 2 different ground truth binarized images of each participant will be accumulated to decide the winner.

D. Participant Methods

Group 1: First, they use two neural network classifiers C_1 and C_2 to classify each pixel as background or not. Two binarized images B_1 and B_2 are got in this step. C_1 is a rough classifier which tries to detect all the foreground pixels while probably making mistakes for some background pixel. C_2 is an accurate classifier which should not classify the background pixel as foreground pixel while probably missing some foreground pixels. Second, they join this two binary images to get the final classification result. In this step, they use each foreground pixel in C_2 as a seed and find all the foreground pixels in B_1 and B_2 which is connected with the seed pixel. All the foreground pixels in B_1 which is not connected to any seed pixel in B_2 are eliminated. The output image in this step is denoted as B_3. Finally, the skeleton image of B_3 is extracted and the dilated skeleton image is the final binary image of the input image. In the training stage, they set the structure of C_1 and C_2 as 124x100x100x1 and 221x100x100x1, respectively. All the transfer functions were set as the sigmoid function. In C_1, the input feature f_i of each pixel P_i is the combination of the RGB values of P_i and the subtractions between the gray value of P_i and the neighborhoods P_j, j=1,...,121. Neighborhoods are the pixels in the 11x11 window with P_i as the window center. In C_2, the feature f_i of each pixel P_i is the combination of the gradient features [5] around P_i and the subtractions between the gray value of P_i and the neighborhoods P_j,j=1,...,121. In the experiment, they used a 45x45 image block around P_i to extract the gradient feature. The image block is divided into 5x5 sub-blocks and 4 direction gradient features are extracted in each sub-block.

Group 2: They employ a Fully Convolutional Network (FCN). It takes a color sub image as input and outputs the probability that each pixel in the sub image is part of the foreground. The FCN is pre-trained on normal handwritten document images with automatically generated "ground truth" binarizations (using the method of Wolf et al [6]). The FCN is then fine-tuned using DIBCO and HDIBCO competition images and their corresponding ground truth binarizations. Finally, the FCN is fine-tuned again on the provided Balinese Palm Leaf images. For inference, the pixel probabilities of foreground are efficiently predicted for the whole image at once and thresholded at 0.5 to create a binarized output image.

Group 3: The submitted method adopts Lu et al.ôs method [7] to estimate the document background based on the polynomial smoothing. The background is estimated by fitting a polynomial function vertically and horizontally. The estimated background is then used to compensate the contrast of the input document image as described in [7]. After that, they apply the local contrast evaluated by the local maximum and minimum [8] to further suppress the background variation on the normalized image. The text stroke edge pixels in the processed image can then be segmented by a global threshold. Finally, the text pixels are classified based on the following criteria: 1) There are text stroke edge pixels existing within a local window of the text pixel, 2) The intensity of the text pixel should be similar to the text stroke edge pixel.

Group 4: They proposed a method named: Hue segmented local contrast binarization. Palm Leaf RGB image is converted into HSV image. Hue is used to segment the foreground and the background from the image. The boundary between the foreground and the background is used as boundary image. The RGB image is combined together to form a gray image using the peak position in the individual histogram as the scaling parameter. The gray image is converted to a local contrast image using Weiner filter. The local contrast image is binarized using Sauvola's threshold N=32,k=0.1, and R=128. The binarized image is filtered using the boundary image and the threshold based on the area, the height, and the width of individual connected component in the image.

E. Results and Analysis

The evaluation results for all submitted methods are presented in Table I. We calculated the average score of each evaluation measure from the 50 images of testing subset with two binary ground truth images. The rank scores for each group are presented in Table II. The method proposed by

Group 2 outperforms all other methods in all evaluation measures.

TABLE I. EVALUATION RESULTS OF CHALLENGE 1

No.	Group	1st GT Binary			2nd GT Binary		
		FM	NRM	PSNR	FM	NRM	PSNR
1.	Group 1	63.32	0.15	31.37	63.24	0.14	31.34
2.	Group 2	68.76	0.13	33.39	68.63	0.12	33.32
3.	Group 3	52.20	0.18	26.92	51.83	0.18	26.83
4.	Group 4	58.57	0.17	29.98	58.32	0.17	29.92

TABLE II. RANK SCORE OF CHALLENGE 1

No.	Group	1st GT Binary			2nd GT Binary		
		FM	NRM	PSNR	FM	NRM	PSNR
1.	Group 1	2nd	2nd	2nd	2nd	2nd	2nd
2.	Group 2	1st	1st	1st	1st	1st	1st
3.	Group 3	4th	4th	4th	4th	4th	4th
4.	Group 4	3rd	3rd	3rd	3rd	3rd	3rd

IV. CHALLENGE 2: QUERY-BY-EXAMPLE WORD SPOTTING ON PALM LEAF MANUSCRIPT IMAGES

A. Description of the Challenge

Word spotting system is one of the most demanding system which has to be developed for the collection of palm leaf manuscript images. This system will facilitate user to find word patch images in all collection of palm leaf manuscript images with a single keyword patch image as a query image. Many image features and descriptors have been proposed to perform the word spotting task. The characteristics of palm leaf manuscripts provide a suitable challenge for testing and evaluation of some image features and descriptors which were already proposed for word spotting methods. Writing in Balinese script, there is no space between words in a text line and some characters are written on upper baseline or under the baseline of text line. The input of this challenge is a set of word-level query patch color images and a set of original color images of palm leaf manuscript. The expected output is a set of spotting area coordinates for each query which are found in the original palm leaf manuscript images.

B. Dataset

To create the word-level annotated ground truth dataset of the manuscript, we asked the Balinese philologists, the students in informatics, and the students in Balinese literature to work together to segment and to annotate the word in manuscript with ALETHEI[1], an advanced document layout and text ground-truthing system [9]. The dataset is partitioned into training and test subsets. For the training subset, we have provided 130 original images and the corresponding 15,022 word-level annotated patch images (Fig. 2). For the testing subset, 100 original images (different from the training subset) and 36 word-level annotated patch images as query test have been provided.

[1] http://www.primaresearch.org/tools/Aletheia

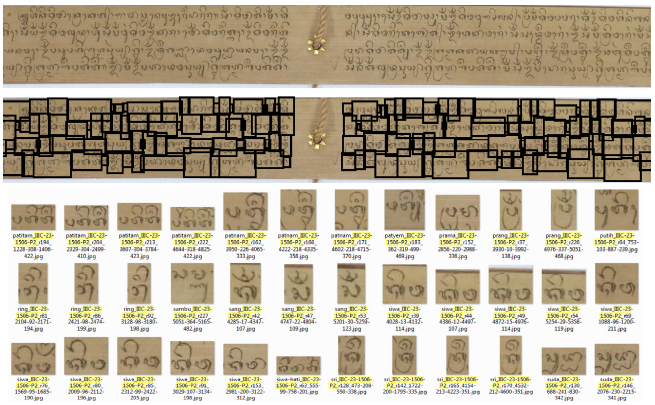


Figure 2. Original image and samples of the word-level annotated patch images

C. Performance Evaluation

We calculated Recall and Precision of spotting area based on ground truth word-level annotated patch images of the testing subset. A spotting area is considered as relevant if it overlaps more than 50% of a ground truth word-level patch area containing the same query word and the size of the spotting area (width and height) is not twice bigger than the size of ground truth area.

Unfortunately, there is no participant who accomplished this challenge. We kindly invite all researchers who work on word spotting method to test their method with our *AMADI_LontarSet*.

V. CHALLENGE 3: ISOLATED CHARACTER RECOGNITION OF BALINESE SCRIPT IN PALM LEAF MANUSCRIPT IMAGES

A. Description of the Challenge

Isolated handwritten character recognition has been the subject of intensive research during the last decades. Balinese script is considered to be one of the complex scripts from Southeast Asia. The alphabet and numeral of Balinese script is composed of ±100 character classes including consonants, vowels, diacritics, and some other special compound characters. The main objective of this challenge is to be able to efficiently and accurately recognize the isolated characters of Balinese scripts. A set of image patches containing individual Balinese character or compound character from the original manuscript will be used as the input, and a correct class of each character should be identified as the output.

B. Dataset

By using our collection of word-level annotated patch images, we applied binarization process to extract automatically all connected component found on the word patch images. The Balinese philologists annotated manually the segment of connected component that represent a correct character in Balinese script. All the data that have been segmented and annotated at character level will serve as our isolated character dataset. The dataset is partitioned into training and test subsets. For the training subset, 11,710 character-level annotated patch images from 130 character classes have been provided (Fig. 3). For the testing subset,

7,673 character-level annotated patch images as query test from 130 character classes have been provided. The number of samples for each classes is different. Some classes are often used in our collection of palm leaf manuscripts, but some others are rarely used.

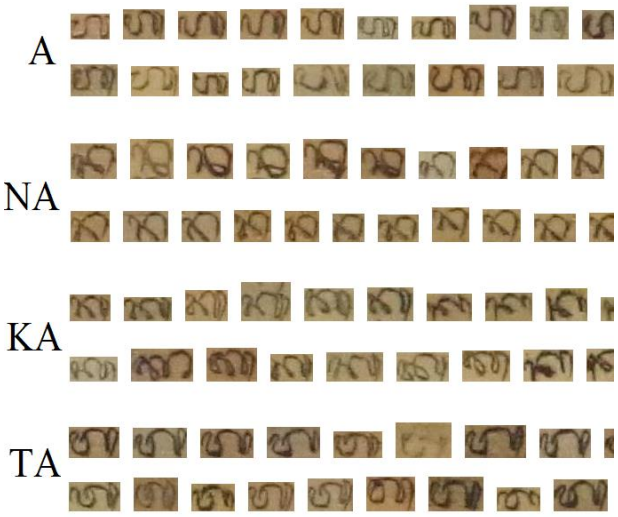


Figure 3. Samples of character-level annotated patch images

C. Performance Evaluation

The results are ranked according to the recognition rate, i.e., the percentage of correctly classified samples over the test samples: C/N, where C is the number of correctly recognized samples, and N is the total number of test samples.

D. Participant Methods

Group 1: (VMQDF and VCMF method) The input image is first processed by the following steps: 1) the input image is resized to 64x64 by linear normalization and median blur with window width 3 is applied on the normalized image. The resulted image is denoted as I_1, 2) the binary image I_b of I_1 is computed by the OTSU method, 3) I_1 is then processed by the pre-processing method described in [5] to overcome the gray scale variation among different images. The mean and scatter employed in the experiment is set as 100 and 100, respectively. The resulted image is denoted as I_n. 97 virtual samples are generated based on I_n by the method proposed in [15]. Gradient feature is extracted using NCGF [5] for each original or virtual sample. MQDF classifier is trained based on the expanded training set. At the test stage, for each input image, 97 virtual images are generated and features are extracted in the same way. VMQDF classifies the input image by the voting of the MQDF results of the input image and all the virtual images. VCMF classifies the input image by the voting of the MQDF+CMF results of the input image and all the virtual images.

Group 3: The submitted method first segments the text pixels out of the background using the adaptive contrast method [8]. To generate more training data, they developed a synthetic data generation method based on the binary images. First, the bounding box of the character can be easily determined in the binary image. They then crop, rotate and warp the origin images without removing the textual information of the origin images. With the synthetic dataset, they build three CNN models to recognize the 133 labels. The CNN model is constructed using the computational network toolkit developed by Microsoft[2]. The input images are resized to 28 by 28. The first CNN network takes the origin images as network input, the second one takes the binary images as network input, and the last one combines the origin and binary images together as network input. Each CNN network has 2 convolution layers followed by two max pooling layers. After training, they further design a random forest to ensemble all the outputs of the CNN networks. The random forest is constructed using Scikit Learn package[3] with 200 trees. The final recognition result is produced by the random forest classifier.

Group 5: The finite impulse response multilayer perceptron (FIRMLP), a class of temporal processing neural networks, is a multilayer perceptron where the weights have been replaced by finite impulse response filters. The FIRMLP consists of only one type of processing unit - the finite impulse response neuron. They have extended this architecture which consists now of layers with different computational properties and recurrent connections as well. The first layer is composed of temporal quadratic units. Hence, those neurons compute polynomial kernels of order two, transforming the input space into a higher dimensional feature space. The second layer uses temporal radial basis function units and separates the extended input space by spheres. The third and fourth layer incorporate sigmoid and linear neurons, respectively, with usual FIR synapses. By using the finite impulse response filter the neuron possess internal short time memory and represents a model for spatio-temporal processing. By adding recurrent connections, the neuron develops long time memory as well. The neural networks used as input 64x64 grey pixel images obtained as follows: 1) first, the MSRCR-Algorithms (MultiScale Retinex with Color Restoration) is applied 2) image binarization is performed by using the K-Means algorithm (from OpenImaj package) with two centers 3) and finally scaling to 64x64 pixel images is applied.

E. Results and Analysis

The evaluation results and ranks for all submitted methods are presented in Table III. The two methods proposed by Group 1 achieved the top two recognition rates. Group 1 obtained a significant 10% higher recognition rate than two other groups. The method VMQDF from Group 1 outperforms all other methods.

TABLE III. EVALUATION RESULTS OF CHALLENGE 3

No.	Group	Recog. Rate (%)	Rank
1.	Group 1: Method VCMF	87.44	2nd
2.	Group 1: Method VMQDF	88.39	1st
3.	Group 3	77.83	3rd
4.	Group 5	77.70	4th

[2] CNTK http://www.cntk.ai/

[3] Scikit-Learn http://scikit-learn.org/

VI. CONCLUSIONS

In this Competition on the Analysis of Handwritten Text in Images of Balinese Palm Leaf Manuscripts, the best result for Challenge 1 (Binarization of Palm Leaf Manuscript Images) was obtained by Chris Tensmeyer, from Computer Science Dept., Brigham Young University. This method achieved the best score for all evaluation measures with two different ground truth images. The best result for Challenge 3 (Isolated Character Recognition of Balinese Script in Palm Leaf Manuscript Images) was obtained by Yunxue Shao, Jiantao Zhou and Guanglai Gao, from College of Computer Science, Inner Mongolia University, with their VMQDF method. This method achieved 88.39% of recognition rate. The dataset of each challenge is available on http://amadi.univ-lr.fr/ICFHR2016_Contest/.

ACKNOWLEDGMENT

The authors would like to thank Museum Gedong Kertya, Museum Bali, and all families in Bali, Indonesia, for providing us the samples of palm leaf manuscripts, and the students from the Department of Informatics Education and the Department of Balinese Literature, Ganesha University of Education for helping us in ground truthing process for this research project. This work is also supported by the DIKTI BPPLN Indonesian Scholarship Program and the STIC Asia Program implemented by the French Ministry of Foreign Affairs and International Development (MAEDI).

REFERENCES

[1] M.W.A. Kesiman, S. Prum, I.M.G. Sunarya, J.-C. Burie, J.-M. Ogier, An Analysis of Ground Truth Binarized Image Variability of Palm Leaf Manuscripts, in: 5th Int. Conf. Image Process. Theory Tools Appl. IPTA 2015, Orleans, France, 2015: pp. 229ï 233.

[2] M.W.A. Kesiman, S. Prum, J.-C. Burie, J.-M. Ogier, An Initial Study On The Construction Of Ground Truth Binarized Images Of Ancient Palm Leaf Manuscripts, in: 13th Int. Conf. Doc. Anal. Recognit. ICDAR, Nancy, France, 2015.

[3] E. Saund, J. Lin, P. Sarkar, PixLabeler: User Interface for Pixel-Level Labeling of Elements in Document Images, in: IEEE, 2009: pp. 646ï 650. doi:10.1109/ICDAR.2009.250.

[4] B. Gatos, K. Ntirogiannis, I. Pratikakis, DIBCO 2009: document image binarization contest, Int. J. Doc. Anal. Recognit. IJDAR. 14 (2011) 35ï 44. doi:10.1007/s10032-010-0115-7.

[5] C.-L. Liu, F. Yin, D.-H. Wang, Q.-F. Wang, Online and offline handwritten Chinese character recognition: Benchmarking on new databases, Pattern Recognit. 46 (2013) 155ï 162. doi:10.1016/j.patcog.2012.06.021.

[6] C. Wolf, J.-M. Jolion, F. Chassaing, Text localization, enhancement and binarization in multimedia documents, in: IEEE Comput. Soc, 2002: pp. 1037ï 1040. doi:10.1109/ICPR.2002.1048482.

[7] S. Lu, B. Su, C.L. Tan, Document image binarization using background estimation and stroke edges, Int. J. Doc. Anal. Recognit. IJDAR. 13 (2010) 303ï 314. doi:10.1007/s10032-010-0130-8.

[8] Bolan Su, Shijian Lu, Chew Lim Tan, Robust Document Image Binarization Technique for Degraded Document Images, IEEE Trans. Image Process. 22 (2013) 1408ï 1417. doi:10.1109/TIP.2012.2231089.

[9] C. Clausner, S. Pletschacher, A. Antonacopoulos, Aletheia - An Advanced Document Layout and Text Ground-Truthing System for Production Environments, in: IEEE, 2011: pp. 48ï 52. doi:10.1109/ICDAR.2011.19.

[10] M. Rusinol, D. Aldavert, R. Toledo, J. Llados, Browsing Heterogeneous Document Collections by a Segmentation-Free Word Spotting Method, in: IEEE, 2011: pp. 63ï 67. doi:10.1109/ICDAR.2011.22.

[11] V. Dovgalecs, A. Burnett, P. Tranouez, S. Nicolas, L. Heutte, Spot It! Finding Words and Patterns in Historical Documents, in: IEEE, 2013: pp. 1039ï 1043. doi:10.1109/ICDAR.2013.208.

[12] T.H. Cormen, T.H. Cormen, eds., Introduction to algorithms, 2nd ed, MIT Press, Cambridge, Mass, 2001.

[13] L. Rothacker, G.A. Fink, P. Banerjee, U. Bhattacharya, B.B. Chaudhuri, Bag-of-features HMMs for segmentation-free Bangla word spotting, in: ACM Press, 2013: p. 1. doi:10.1145/2505377.2505384.

[14] L. Rothacker, M. Rusinol, G.A. Fink, Bag-of-Features HMMs for Segmentation-Free Word Spotting in Handwritten Documents, in: IEEE, 2013: pp. 1305ï 1309. doi:10.1109/ICDAR.2013.264.

[15] Y. Shao, C. Wang, B. Xiao, Fast self-generation voting for handwritten Chinese character recognition, Int. J. Doc. Anal. Recognit. IJDAR. 16 (2013) 413ï 424. doi:10.1007/s10032-012-0194-8.