ICDAR2015 Competition on Signature Verification and Writer Identification for On- and Off-line Skilled Forgeries (SigWIcomp2015)

Muhammad Imran Malik*, Sheraz Ahmed*, Angelo Marcelli†,
Umapada Pal‡, Michael Blumenstein§, Linda Alewijns¶, Marcus Liwicki‖
*German Research Center for AI (DFKI GmbH), Kaisersautern, Germany (firstname.lastname@dfki.de)
†University of Salerno Via Giovanni Paolo II, Fisciano (SA), Italy (amarcelli@unisa.it)
‡CVPR Unit, Indian Statistical Institute, Kolkata, India (umapada@isical.ac.in)
§School of ICT, Griffith University, Australia (m.blumenstein@griffith.edu.au)
¶Netherlands Forensic Institute. The Hague, The Netherlands. (l.alewijnse@nfi.minvenj.nl)
‖University of Fribourg, Switzerland (marcus.liwicki@unifr.ch)

***Abstract*—This paper presents the results of the ICDAR 2015 competition on signature verification and writer identification for on- and off-line skilled forgeries jointly organized by PR-researchers and Forensic Handwriting Examiners (FHEs). The aim is to bridge the gap between recent technological developments and forensic casework. Two modalities (signatures and handwritten text) are considered and training and evaluation data are collected and provided by FHEs and PR-researchers. Four tasks are defined for four different languages; Bengali off-line signature verification, Italian off-line signature verification, German on-line signature verification, and English handwritten text based writer identification. In total, 40 systems have participated in this competition. The participants of the signatures modality were motivated to report their results in Likelihood Ratios (LRs). This has made the systems even more interesting for application in forensic casework. For evaluating the performance of the systems, we have used the forensically substantial Cost of Log Likelihood Ratios ($\widehat{C}_{llr}$) in the case of signatures, and the F-measure in the case of handwritten text.**

I. Introduction

Most of the current research in the field of automatic signature verification does not take the real needs of Forensic Handwriting Experts (FHEs) into account [1], [3], [4], [5]. In their casework, FHEs often need to have verification results in the form of likelihood ratios. This is very important as it allows one to combine the FHE's evidence (from the results of an automated system) with other evidence presented in the courts of law [6], [7]. We have organized the ICDAR SigWIcomp2015 competition where we asked the participants to produce a comparison score (e.g., a degree of similarity or difference), and the evidential value of that score, expressed as the ratio of the probabilities of finding that score when the questioned signature is a genuine signature and when it is a forgery (i.e., the likelihood ratio). Note that by these competitions we are further strengthening a paradigm shift in automatic signature verification (introduced in the SigComp 2011, and tested heavily in 4NSigComp 2012 and SigWiComp2013) from the "decision paradigm" to an "evidential value paradigm". For more details on this, please refer to [8], [9], [10].

TABLE I
Italian Off-line signature data

Data	Authors	Genuine	Forged	Total
Training	50	250	0	250
Testing	50	229	249	478

TABLE II
Bengali Off-line signature data

Data	Authors	Genuine	Forged	Total
Training	10	120	0	120
Testing	10	120	300	420

Furthermore, we have focused on the writer identification and retrieval task from the view point of FHEs. Here the defined task is: *Given a handwritten text from an author, retrieve all the texts written in different writing styles by the same author*. This is also very important when viewed from the perspective of FHEs. They often have the scenarios where they are provided with a piece of text handwritten by an author in a different style than those texts which are present in the dataset and are also written by the same author. This is a usual situation when the same person, for example, writes threat letters to more than one persons and tries to change her/his handwriting style every time. The task here is to link the different handwriting styles from the same writer; which in fact is a real challenge so far. For evaluation of this task, we used standard precision and recall measures and calculated F-measure to declare the winning system.

II. Data

We have the signature data available in Bengali, Italian, and German and handwriting data in different writing styles in English only. The details follow.

A. *Italian Off-line Signatures*

The Italian off-line signatures were collected at the University of Salerno from the University employees and students.

TABLE III
GERMAN ON-LINE SIGNATURE DATA

Data	Authors	Genuine	Forged	Total
Training	30	300	0	300
Testing	30	150	300	450

TABLE IV
ENGLISH OFF-LINE HANDWRITTEN TEXT DATA

Data	Authors	Pages/Author	Total
Training	55	3	165
Testing	55	3	165

Two unique aspects of these data are; these are actual signatures that were provided by individuals while filling various forms and applications at the University, and the signatures collection span over a period of time that is between 3 and 5 years depending on the subject. The forgeries are also made by students where they were allowed to practice the forgery as many times as they liked and then they produced skilled forgeries. For training, a set of 50 specimen genuine authors, with 5 reference signatures from each specimen author, was provided to the participants in binarized form. For evaluation, we used the same 50 specimen authors; with 10 corresponding questioned signatures each: containing genuine signatures and skilled forgeries, in binarized form. Detailed breakup is provided in Table I.

B. Bengali Off-line Signatures

The signatures were collected from different parts of the West Bengal, a state of India. The majority of the signatures were contributed by students. A total number of 240 genuine signatures were collected from 10 contributor (24 genuine signatures by each). For each contributor, all genuine specimens were collected in a single day's writing session. In order to produce the forgeries, the imitators were allowed to practice forgeries as long as they wished. A total number of 300 (30 signatures, 10 individuals) forged signatures were collected. The images were captured in 256 level Grey scale at 300 dpi and stored in TIFF format (Tagged Image File Format). A detailed breakup is provided in Table II.

C. German On-line Signatures

The German on-line signatures were collected at the German Research center for Artificial Intelligence, Germany. A unique aspect of these data is that the data have been collected using a digitized pen rather than a tablet, i.e., by Anoto Pen [11]. This pen specializes in providing the look and feel of regular pens. It only demands to add Anoto dot pattern to any paper and data can be digitized seamlessly. The Anoto pattern makes it possible for the Anoto pen built-in camera to detect strokes and record signatures. The signature data were collected from employees of different financial institutions and students of University of Kaiserslautern who also participated in generation of skilled forgeries. The dataset consists of ASCII files with the format: X, Y, and Pressure with sampling rate of 75 Hz and resolution of 85 dpi. For training, 30 specimen genuine authors were selected, from whom each one provided 10 genuine reference signatures. For evaluation, data from the same 30 specimen genuine authors; with 15 corresponding questioned signatures each: containing genuine signatures and skilled forgeries were used. Detailed breakup is provided in Table III.

Note that for all of the signature verification tasks, we did not provide any forgeries for training (see Tables I, II, and III). This is important as; In the real world, one can never limit the forgery set since every signature other than the genuine signature written naturally by an authentic author can be a forgery, and also when forgeries are used for training–there is always a chance that an automatic system may learn to declare signatures as forgeries when they are coming from the forgers on whom the system is trained [12], [13].

D. English Handwritten Text Samples

The handwritten text samples for the writer identification (based on different handwriting styles) task were off-line and in English only. They were collected and prepared by the Netherlands Forensic Institute for this competition/task. A pre-printed paper was used with horizontal lines for writing. The pre-printed paper was placed underneath the blank writing paper. Four extra blank pages were added underneath the first two pages to obtain a soft writing surface. The writings were scanned at 400 dpi using the EPSON Expression 10000XL scanner, with RGB color and saved as PNG images. Files were randomly numbered, so that file numbers do not link to the writers. The training set comprised of 165 samples written by all the 55 authors in 3 different handwriting styles. The evaluation set comprised of 165 samples written by the same 55 authors in 3 different handwriting styles (other than the styles available in the training set). A detailed breakdown is provided in Table IV.

III. SUBMITTED SYSTEMS

50 systems initially registered for this competition, however, 40 systems from 13 different institutions eventually participated. The details follow: 9 systems for off-line Italian signature verification, the same 9 systems optimized for off-line Bengali signature verification, 12 system for the on-line German signature verification, and 10 systems for the writer identification task. Table V provides details about affiliations of the participants. In the following we will describe each of these systems briefly by providing references so that interested readers may follow. Note that some of the participants preferred remaining anonymous after the results were declared.

A. Off-line Verification Tasks: Italian and Bengali

System 1: This system uses histogram of oriented gradients (HOG) and local binary patterns (LBP) extracted from local regions, together with the user-based and global SVM classifiers and a sophisticated classifier combination. The basic system description can be found in [14].

TABLE V
OVERVIEW OF THE SUBMITTED SYSTEMS

System	Modality	Participant	Mode
1	Signatures	Sabanci University, Turkey	Off-line
2-6	Signatures	Tebessa University, Algeria	Off-line
7,8	Signatures	Qatar University, Qatar	Off-line
9	Signatures	Commercial System	Off-line
10-14	Signatures	Anonymous	On-line
15	Signatures	Anonymous	On-line
16	Signatures	Bahria University, Pakistan	On-line
17-18	Signatures	Cursor Insight	On-line
19	Signatures	Commercial System	On-line
20	Signatures	Qatar University, Qatar	On-line
21	Signatures	Sabanci University, Turkey	On-line
22-26	Handwriting	Tebessa University, Algeria	Off-line
27-31	Handwriting	Qatar University, Qatar	Off-line

TABLE VI
RESULTS FOR TASK 1: ITALIAN OFF-LINE SIGNATURE VERIFICATION

System	Participant	$\widehat{C}_{llr}$	$\widehat{C}_{llr}^{min}$
1	**Sabanci University**	**0.655109**	**0.021358**
2	Tebessa University	0.993189	0.893270
3	Tebessa University	1.065696	0.952499
4	Tebessa University	1.074474	0.880930
5	Tebessa University	1.065475	0.901003
6	Tebessa University	1.041895	0.901003
7	Qatar University	8.901864	0.972708
8	Qatar University	13.111064	0.960163
9	Commercial System	1.003786	0.988845

System 2: This system is based on edge directional based features. The method starts with conventional edge detection that generates a binary image in which only the edge pixels are "on". Then each edge pixel is considered in the middle of a square neighborhood and checked in all directions emerging from the central pixel and ending on the periphery of the neighborhood for the presence of an entire edge fragment. For comparison purposes, the Manhattan Distance Metric is used.

System 3: This system is based on the edge-hinge features which estimate the joint distribution of edge angles in a writer's handwriting. They are constructed by performing an edge detection using a Sobel kernel on the input images, and subsequently, measuring the angles of both edge segments that emanate from each edge pixel.

System 4: The system is based on multi-scale run length features which are determined on the binary image taking into consideration both the black pixels corresponding to the ink trace and the white pixels corresponding to the background. For details, refer to [15].

System 5: The system is based on the combination of both types of features used by the previous two methods: multi-scale edge-hinge features and multi-scale run-length features [16].

System 6: This system is based on the combination of three types of features used by the systems 2, 3, and 4: multi-scale edge-hinge features, multi-scale run-length features and the edge based directional features.

System 7 and System 8: Both of these system combine through a logistic regression classifier the geometrical features

TABLE VII
RESULTS FOR TASK 2: BENGALI OFF-LINE SIGNATURE VERIFICATION

System	Participant	$\widehat{C}_{llr}$	$\widehat{C}_{llr}^{min}$
1	Sabanci University	0.68895	0.052485
2	Tebessa University	0.933750	0.154390
3	Tebessa University	0.939850	0.116541
4	**Tebessa University**	**0.923886**	**0.039721**
5	Tebessa University	0.931253	0.060248
6	Tebessa University	0.929922	0.055556
7	Qatar University	1.161387	0.973043
8	Qatar University	2.839290	0.297842
9	Commercial System	0.998630	0.893893

described in [17]. Those features are based on number of holes, moments, projections, distributions, position of barycenter, number of branches in the skeleton, Fourier descriptors, tortuosities, directions, curvatures, chain codes and edge based directional features. The two systems differ in weighting the features they are built on.

System 9: This is a commercial system and the participating organization has requested to not mention their identity. The following basic details are provided about this system. The system has four main modules. The first one identify and extracts the signature from the image page. The second module applies filtering and other preprocessing on the extracted signature. The third module extracts various spatial, geometric, morphological, and statistical features. The last module allows the combination of multiple classifiers and uses DTW, Pearson Correlation, and Euclidean distance combined by an MLP neural network to make the final decision.

The same nine systems (Systems 1-9) have participated for both Italian and Bengali tasks (when optimized for the specific task), their details are therefore omitted.

B. *On-line Verification Task: German*

System 10-14: Systems 10 to 14 are from the same academic organization and they have requested to not mention their identity. All of these systems are based on DTW on various time functions derived from the position trajectory and the pressure information. These 5 systems differ in the specific functions considered and the type of score normalization applied.

System 15: The following details are available about the working of this system, though the participating organization has chosen to remain anonymous. This system is based on the DTW algorithm and uses both global features (e.g., total writing time) and local features (e.g., pressure in a given point). If the total writing time of the signature to be verified is acceptable, for each local feature f (such as X-Y coordinates, pressure), the DTW distance between the time series related to the authentic signature and to the signature to be verified is computed. Then, DTW distances are combined with a weighted sum, by giving a weight to each of the different kinds of features. The similarity score is computed accordingly.

System 16: In this system, multiple time series representation of signature including time series of X-coordinate, Y-Coordinate, Pressure, speed and centroid distance are com-

TABLE VIII
RESULTS FOR TASK 3: GERMAN ON-LINE SIGNATURE VERIFICATION

System	Participant	$\widehat{C}_{llr}$	$\widehat{C}_{llr}^{min}$
10	Anonymous	0.948878	0.680014
11	Anonymous	0.944402	0.688105
12	Anonymous	0.855595	0.576762
13	Anonymous	0.874655	0.577098
14	Anonymous	1.262207	0.684602
15	Anonymous	0.996586	0.839468
16	Bahria University	1.025046	0.772073
17	**Cursor Insight**	**1.243891**	**0.290158**
18	Cursor Insight	1.496257	0.540563
19	Commercial System	0.807276	0.465681
20	Qatar University	3.345016	0.854712
21	Sabanci University	0.873383	0.304159

TABLE IX
RESULTS FOR TASK 4: WRITER IDENTIFICATION BASED ON DIFFERENT HANDWRITING STYLES.

System	Participant	Avg. F1-measure(%)
22	Tebessa University	17.37
23	Tebessa University	33.54
24	Tebessa University	32.53
25	**Tebessa University**	**33.94**
26	Tebessa University	33.54
27	Qatar University	30.71
28	Qatar University	21.01
29	Qatar University	21.01
30	Qatar University	20.80
31	Qatar University	19.60

puted. Verification of genuine and forged signature is done using adaptation of multivariate m_Mediods based classification and anomaly detection approach as presented in [18].

System 17-18: Both of these systems are commercial products designed by Cursor Insight. These systems specialize in calculating more than 70000 movement characteristics of each handwritten sample. For further details, refer to Cursor Insight [1].

System 19: This system compares the signatures using statistical data as well as a time based model. It extracts the changes of several dynamic features (direction, pressure and speed) over time. The statistical data contains static information like duration of the signature or duration and time of pen liftings. The reference signatures are compared against each other to define the validity space of the model. The similarities of the features are combined into a final score using different weightings.

System 20: This system computes differences between the features of the questioned signature and the reference signatures at the signal level as well as the histogram level [19]. Those features include x and y coordinates, pressure, directions, angles, speed and angular speed.

System 21: This system only uses the 'x' and 'y' coordinates as features and Dynamic Time Warping (DTW) as the matching algorithm. Normalization of the query scores are done based on reference signature statistics. Details of this system can be found in [20].

C. Off-line Writer Identification Task: English

Systems 22-26: By the same participant (who provided Systems 2-6), with the same methodology, as explained earlier, optimized to consider handwriting.

Systems 27-31: The proposed methods combine through a logistic regression classifier the geometrical features. Those features are based on tortuosities, directions, curvatures, chain codes and edge based directional features. Further details are available in [17], [19].

IV. EXPERIMENTS AND EVALUATION

We defined two modalities, i.e., signatures and handwritten text, and the following four tasks for this competition,

- Task 1: Italian off-line signature verification
- Task 2: Bengali off-line signature verification
- Task 3: German on-line signature verification
- Task 4: Writer identification based on handwriting styles

Accordingly, the evaluations are made separately for each of these tasks. In the case of signature verification tasks (the first three: both on-line and off-line), the task was to determine if a given questioned signature has been written by the author of the n reference specimen signatures or if it was forged by another writer. We evaluated the signature verification systems according to the cost of the log-likelihood ratios $\widehat{C}_{llr}$ using the FoCal toolkit [21], and finally the minimal possible value of $\widehat{C}_{llr}$, i.e., $\widehat{C}_{llr}^{min}$, as the final assessment value. Note that a smaller value of $\widehat{C}_{llr}^{min}$ denotes a better performance of the method. An interesting observation about the $\widehat{C}_{llr}$ and $\widehat{C}_{llr}^{min}$ is that it is not only based on the number of errors made by a system but also looks into the severity of errors by warping the scores of automatic systems. This makes the metric ($\widehat{C}_{llr}^{min}$) well suited for forensic applications where a severely mistaken system (although may be having a very low error rate) may lead a person to death, therefore considering the severity of errors is substantial for forensic casework. We have focused on this issue in the previous competitions of this chain, i.e., at ICDAR 2011 and ICDAR 2103. For more details about this, refer to [10].

Further important observations about the signature verification tasks for different data are: For Italian off-line signatures, the system using a combination of local and global features outperformed all other systems (mostly based on global features) by a large margin; for Bengali off-line signatures, most of the systems performed good and there was not much variance among the results which might indicate that the systems which were originally designed for western languages were able to adapt well on an Indic language provided better training; for German on-line signatures, the systems had to face difficulties as these data were collected by specialized electronic pen-Anoto. The Anoto pen also records position and pressure similar to standard data capturing tablets, however with different internal data capturing mechanism. Since, the use of electronic pens is increasing, the PR research should also consider these data and optimize the systems accordingly.

[1] www.cursorinsight.com

The results of the first three tasks are given in Tables VI, VII, and VIII, respectively. As can be seen, different systems performed better on different tasks. The winner of the Italian off-line signature verification task, Task 1, is system 1 from Sabanci University, Turkey. The winner of the Bengali off-line signature verification task, Task 2, is system 4 from Tebessa University, Algeria. The winner of the German on-line signature verification task, Task 3, is system 17 from Cursor Insight (a company).

The results of the fourth task, i.e., writer identification by linking various handwriting styles of the same writer are provided in Table IX. In this task we used the standard precision and recall measures to evaluate the systems and final ranking is based on the best average F-measure value [22]. Note that in this case the values of precision and recall were the same (and thereby the F-measure value–so only F-measure value is given): we used the F1-measure, the harmonic mean of precision and recall, where equal emphasis is given to both the precision and recall. This is due to the scenario for the said competition where for every author 3 documents handwritten in different styles were available in the repository of all the documents and the system must also retrieve only the 3 top matching handwritten documents from that repository.

Furthermore, as given in Table IX, the systems are quite low in performance on the F-measure scale. This represents that the given task was really challenging and still requires a lot of efforts from the PR-researchers. Nonetheless, System 25 performed comparatively better than the other systems and is declared the winner for this task.

V. Conclusions and Outlook

In this paper we presented the results of the ICDAR2015 competition on signature verification and writer identification for on- and off-line skilled forgeries (SigWiComp2015) jointly organized by PR-researchers and FHEs. Note that it is a continuation of previous signature verification competitions, i.e., SigComp2011, 4NSigComp2012, and SigWiComp2013.

In the SigWiComp2015, we have defined three tasks for off-line and on-line signature verification and provided signature data (both on- and off-line) in multiple languages (Italian, Bengali, and German). Furthermore, for evaluation, we reinforced our paradigm based on likelihood computations via the $\widehat{C}_{llr}$ and $\widehat{C}_{llr}^{min}$ values which are forensically substantial. In addition to that, in the SigWiComp2015, we also defined one task for writer identification and retrieval based on different handwriting styles. The inclusion of this task has made the SigWiComp2015 even more forensically relevant and interesting for both the FHEs and PR-researcher.

In the future, we plan to make the datasets even larger and more diverse. We also plan to include signature samples written in different languages (bi-lingual signatures). Another important aspect is the usability and application of automatic systems in real forensic cases, which motivates us to evaluate the usability of automatic systems as well. We are planning to organize competitions focusing on the real world usability of automatic verification systems in the future.

VI. Acknowledgement

We are thankful to Antonio Parziale and Srikanta Pal who provided us access to the signature data for SigWIcomp2015.

References

[1] R. Plamondon and G. Lorette, "Automatic signature verification and writer identification – the state of the art," *Pattern Recognition*, vol. 22, pp. 107–131, 1989.

[2] F. Leclerc and R. Plamondon, "Automatic signature verification: the state of the art 1989–1993," in *Progress in Automatic Signature Verification*, R. Plamondon, Ed. World Scientific Publ. Co., 1994, pp. 13–19.

[3] R. Plamondon and S. N. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63–84, 2000.

[4] D. Impedovo and G. Pirlo, "Automatic signature verification: The state of the art," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 38, no. 5, pp. 609–635, Sep. 2008.

[5] S. Ahmed, M. I. Malik, M. Liwicki, and A. Dengel, "Signature segmentation from document images," in *ICFHR*. IEEE, 2012, 423-427.

[6] J. Gonzalez-Rodriguez, J. Fierrez-Aguilar, D. Ramos-Castro, and J. Ortega-Garcia, "Bayesian analysis of fingerprint, face and signature evidences with automatic biometric systems," *Forensic Science International*, vol. 155, no. 2-3, pp. 126–140, 2005.

[7] M. I. Malik and M. Liwicki, "From terminology to evaluation: Performance assessment of automatic signature verification systems," in *ICFHR*. IEEE, 2012, pp. 613–618.

[8] M. Liwicki, M. I. Malik, C. E. van den Heuvel, X. Chen, C. Berger, R. Stoel, M. Blumenstein, and B. Found, "Signature verification competition for online and offline skilled forgeries SigComp2011," in *ICDAR*. IEEE, 2011, pp. 1480–1484.

[9] M. Liwicki, M. I. Malik, L. Alewijnse, C. E. van den Heuvel, and B. Found, "Icfhr2012 competition on automatic forensic signature verification 4NsigComp 2012," in *ICFHR*. IEEE, 2012, pp. 780–785.

[10] M. I. Malik, M. Liwicki, L. Alewijnse, W. Ohyama, M. Blumenstein, and B. Found, "ICDAR 2013 competitions on signature verification and writer identification for on-and offline skilled forgeries (SigWiComp 2013)," in *ICDAR*. IEEE, 2013, pp. 1477–1483.

[11] M. I. Malik, S. Ahmed, A. Dengel, and M. Liwicki, "A signature verification framework for digital pen applications," in *DAS*. IEEE, 2012, pp. 419–423.

[12] M. I. Malik, L. Alewijnse, M. Liwicki, and M. Blumenstein, "Signature verification tutorial," *2nd International Workshop and Tutorial on Automated Forensic Handwriting Analysis (AFHA)*, 2013.

[13] M. Liwicki and M. I. Malik, "Surprising? power of local features for automated signature verification," in *15th IGS Conf.* International Graphonomics Society, 2011, pp. 18–21.

[14] M. B. Yilmaz, B. Yanikoglu, C. Tirkaz, and A. Kholmatov, "Offline signature verification using classifier combination of hog and lbp features," in *Int. Joint Conf. on Biometrics*, 2011, pp. 1–7.

[15] C. Djeddi, I. Siddiqi, L. Souici-Meslati, and A. Ennaji, "Text-independent writer recognition using multi-script handwritten texts," *Pattern Recognition Letters*, vol. In press, 2013.

[16] C. Djeddi, L. Souici-Meslati, and A. Ennaji, "Writer recognition on arabic handwritten documents," in *ICISP*, 2012, pp. 493–501.

[17] A. Hassaine, S. Al-Maadeed, and A. Bouridane, "A set of geometrical features for writer identification," *Neural Information Processing. Springer Berlin Heidelberg*, 2012.

[18] S. Khalid and S. Razzaq, "Frameworks for multivariate m-mediods based modeling and classification in euclidean and general feature spaces," *Pattern Recogn.*, vol. 45, no. 3, pp. 1092–1103, 2012.

[19] A. Hassane and S. Al-Maadeed, "An online signature verification system for forgery and disguise detection," *Neural Information Processing. Springer Berlin Heidelberg*, 2012.

[20] A. Kholmatov and B. Yanikoglu, "Identity authentication using improved online signature verification method," *Pattern Recognition Letters*, vol. 26, no. 15, pp. 2400–2408, 2005.

[21] N. Brümmer and J. du Preez, "Application-independent evaluation of speaker detection," *Computer Speech & Language*, vol. 20, no. 2-3, pp. 230–275, 2006.

[22] T. Fawcett, "An introduction to ROC analysis," *Pattern Recogn. Lett.*, vol. 27, pp. 861–874, June 2006.