Face demorphing in the presence of facial appearance variations

Matteo Ferrara, Annalisa Franco, Davide Maltoni
Department of Computer Science and Engineering
University of Bologna
Via Sacchi, 3- 47521 Cesena (Italy)
{matteo.ferrara, annalisa.franco, davide.maltoni}@unibo.it

***Abstract*— This study focuses on the robustness of face demorphing as a technique to protect face recognition systems against the well-known morphing threat. In particular, we check if in presence of face variations of different type and strength, the demorphing process significantly reduces the Genuine Acceptance Rate leading to an excessive number of false morphing warnings. Experimental results show that, except for extreme conditions that are unlikely in e-gates scenario, demorphing does not markedly affect face recognition accuracy.**

Keywords*— *Automated border control, eMRTD, face demorphing, face morphing attack, face recognition, face variations.

I. INTRODUCTION

Recent works [*1*] [*2*] confirmed the feasibility of the morphing attack in the context of Machine Readable Travel Documents (eMRTD). At today, this remains one of the major security threat for Automated Border Control Systems (ABC). In fact, ABC can easily be fooled by morphed face images containing combined facial features of two different subjects, which can share the same document to travel illicitly. In particular, if a morphed image, which is similar enough to the face of two subjects, can be inserted in an eMRTD, a criminal could exploit the passport of an accomplice with no criminal records to overcome the security controls.

A promising solution for morphing detection – Face Demorphing – has been recently proposed in [*3*] to detect the morphing attack. The approach is based on the idea of detecting anomalies at the gate where the live image of the passenger can be compared against the eMRTD one; in the demorphing approach a (potentially) morphed image stored into the document is reverted (or demorphed) to reveal the identity of the legitimate document owner, thus allowing the system to issue a warning.

Fig. 1 shows the complete workflow of the face verification system at the ABC gate, in the hypothesis of including a morphing attack detection module based on face demorphing. If the passenger passes the standard face verification stage, the demorphing step takes place: the document face image is demorphed and compared to the live image; in case of a negative match a "morphing" warning is issued to a second line officer. Of course, to be useful in practice, both the morphing acceptance rate and the amount of false morphing warnings must be sufficiently small.

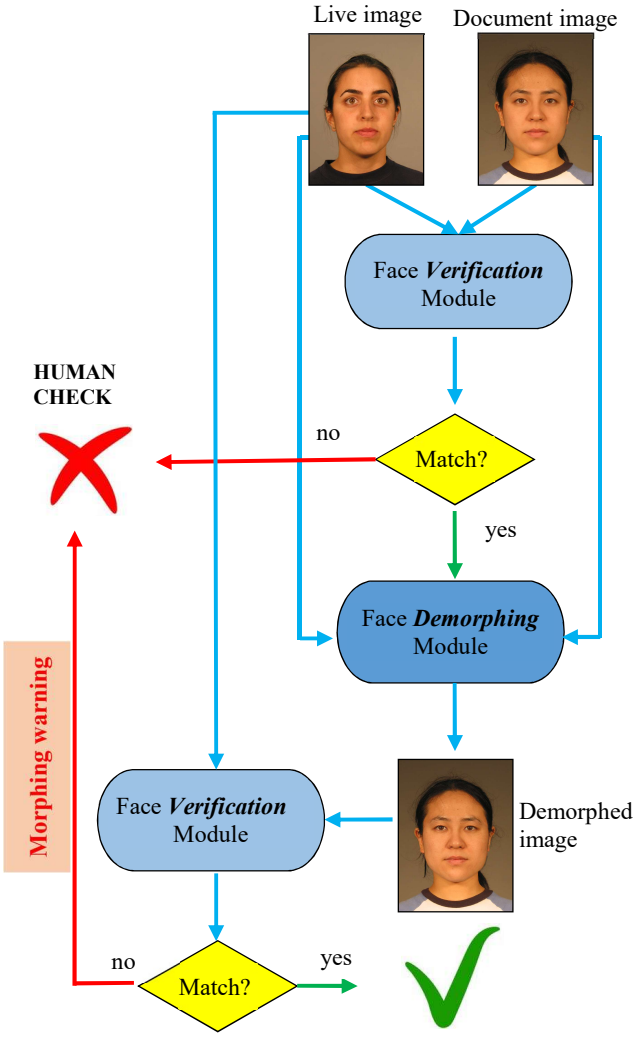

Fig. 1 Functional schema of the face verification procedure performed at ABC gates, including a face demorphing module to spot possible morphing attacks.

The experiments carried out in [*3*] show that face demorphing can substantially reduce the success rate of the morphing attack to ABC systems and, at the same time, keeps the number of false morphing warnings quite limited. This is a fundamental requirement for the practical applicability of the morphing detection techniques since an increment of the false alarms would significantly affect the normal functionalities of the gate. However in the datasets used in [*3*] the face pose is frontal, lighting is also frontal and expression is neutral.

ISBN 978-90-827970-1-5 © EURASIP 2018

The aim of this work is to evaluate the effects of face demorphing on the genuine images (not morphed) in the presence of facial appearance variations. Even though the eMRTD context represents a quite controlled scenario, some face image variations are still possible:

1) image manipulations applied to the digital image during document issuance (i.e., printing/scanning of the ID photo, image compression);

2) typical facial appearance variations observed at the gate in a real operational scenario (e.g., pose and lightning changes, presence of accessories, non-neutral facial expressions).

Some studies and recent evaluation campaigns, focused on the analysis of the effects of such variations on face recognition accuracy (see [*4*] [*5*] [*6*]). In general, the literature results show that such factors represent the major challenges for automatic face recognition. For this reason the effects of the above mentioned alterations on the face demorphing accuracy have to be carefully evaluated since they could determine a significant increment of the false morphing warnings (i.e., a reduction of the genuine acceptance rate), thus reducing the effectiveness of demorphing as a protection of ABC.

II. Face Morphing and Demorphing

Face morphing is an operation that transforms a face image into another, based on a set of reference points. It consists of a combination of geometric warping, aimed at aligning the reference points of the two faces, and image blending which produces a smooth average image starting from the two originals. Such process can be ideally inverted to nullifying the effects of a possible morphing operated on a document image. This is the idea behind face demorphing, recently proposed in [*3*] as a possible countermeasure to the morphing attack. For lack of space the details of morphing/demorphing process cannot be described here; interested readers can refer to [*3*] for more information. It is worth noting that in a real scenario, the demorphing process is not exact since the image acquired live at the gate does not coincide with that used to create the morphed image. The typical face variations characterizing face images is the context of eMRTD could amplify such approximation and lead to many false morphing warnings.

III. Databases and Experiments

The experiments have been carried out using four commercial face recognition SDKs (referred to as $Sdk_i, i = 1,..,4$) which provided top performance in the recent Face Recognition Vendor Test (FRVT) Ongoing [*7*] [*8*]; the names of the SDKs cannot be disclosed and the results will be therefore presented in anonymous form. All the SDKs fulfill the operational conditions suggested by Frontex for ABC gates (a maximum False Rejection Rate of 5% at a False Acceptance Rate of 0.1%) [*9*]. While in [*3*] both genuine and morph attempts have been tested, in this paper we focus on genuine comparisons only. Therefore, the performance is here evaluated in terms of Genuine Acceptance Rate (GAR), also known as True Match Rate (TMR), with the aim to quantify the accuracy drop determined by demorphing. To this purpose we measure the GAR for the different SDKs before (reference performance) and after applying face demorphing. According to the outcomes in [*3*], a fixed demorphing factor of 0.25 is here used. The security thresholds indicated by the SDK documentations to achieve FAR of 0.1% is set for each SDK, to make them working at the same operating point.

A. Variations of the document image

The first set of experiments focuses on the two main variations applied to the ID photo to be included into a document: printing/scanning and image compression. The MorphDB [*3*] is used for these tests; MorphDB was specifically designed to reproduce the scenario where the ID photo is printed on photographic paper and then scanned by an officer. Moreover a JPEG2000 compression [*10*] is applied at different rates, both to the digital and printed/scanned images to emulate the whole document issuing process. The compression is performed by fixing the size (Kb) of the compressed image.

TABLE I - SDK1 JPEG2000 COMPRESSION: COLUMNS BDE AND ADE REPORT GAR PERCENTAGE BEFORE AND AFTER THE DEMORPHING PROCESS, RESPECTIVELY.

Format	Size (Kb)	BDe	ADe
Digital	4	99.9	99.2
	6	100.0	99.6
	8	100.0	100.0
	10	100.0	99.9
	15	100.0	100.0
	18	100.0	100.0
	20	100.0	100.0
	25	100.0	100.0
	No compression	100.0	100.0
P&S	4	100.0	99.3
	6	100.0	99.7
	8	100.0	100.0
	10	100.0	99.9
	15	100.0	100.0
	18	100.0	100.0
	20	100.0	99.9
	25	100.0	100.0
	No compression	100.0	100.0

Table I reports only the results of Sdk1: small drops can be sometimes observed after the demorphing process. This is not the case of other three SDKs where a 100% GAR is always obtained. These results show that (i) both the printing/scanning process and the image compression (even at low image size) do not affect the recognition performance and (ii) the demorphing process does not affect the accuracy, in particular for image size of 10Kb or higher as suggested in the ISO standard [*11*].

B. Variations of the live image

The evaluation of the possible variations of the live image acquired at the gate has been conducted on the CAS-PEAL-R1 dataset [*12*] containing accessory, expression, lighting, and pose changes. Throughout these tests, a neutral image is compared against an image with a specific variation.

Fig. 2 reports the results of the different SDKs on the *Accessory* testing set (2616 comparisons), whose subjects wear different glasses (A1-A3) and hats (A4-A6). Overall, the SDKs

ISBN 978-90-827970-1-5 © EURASIP 2018

exhibit a good robustness to the presence of accessories; an exception is represented by A3 corresponding, in many cases, to sunglasses covering the eyes region. Face demorphing preserves in general a good accuracy, with the only exception of the test Sdk3 -A3 where a certain performance drop is observed. The quality of face demorphing in the presence of sunglasses is reduced for two main reasons: i) the landmarks in the eyes region cannot be located; ii) the texture of the eyes region is completely different and introduces some small artifacts to which Sdk3 is particularly sensitive. However, according to the Frontex best practices for ABC gates, an image quality check is recommended to assure that face and eyes can be clearly located and the use of sunglasses would not be accepted at gate.

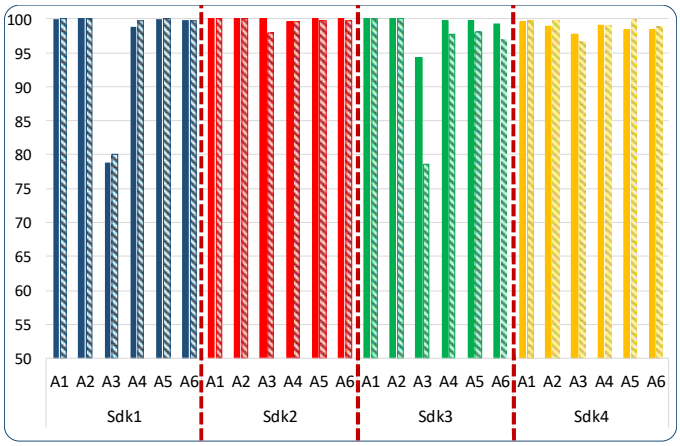

Fig. 2 GAR of the four SDKs (Sdk1: blue, Sdk2: red, Sdk3: green, Sdk4: yellow) for the *Accessory* testing set of the CAS-PEAL-R1 database, before (solid bar) and after demorphing (dashed bar). A1, A2 and A3 correspond to glasses, A4, A5 and A6 to hats.

The effects of unnatural facial expressions have been evaluated performing 1884 comparisons. The expressions represented in the dataset are: eyes closed (EC), frowning (EF), laughing (EL), mouth open (EO), surprising (ES). The results are reported in Fig. 3 for the different SDKs. Also in this case the SDKs show a good robustness and ability to deal with such facial variations. The application of face demorphing does not affect the recognition accuracy.

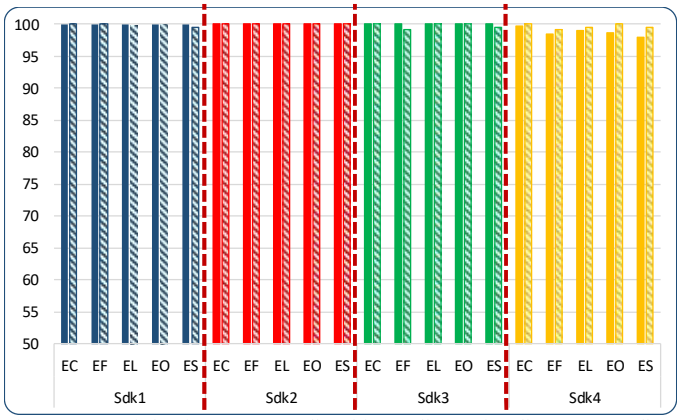

Fig. 3 GAR of the four SDKs (Sdk1: blue, Sdk2: red, Sdk3: green, Sdk4: yellow) for the *Expression* testing set of the CAS-PEAL-R1 database, before (solid bar) and after demorphing (dashed bar).

The effects of lighting changes are described in Fig. 4. The *Lighting* test consists of 1615 comparisons where the position of the light source (fluorescent light) is systematically varied in terms of azimuth (-45°, 0°, +45°) and elevation (U: up, M: middle, D: down).

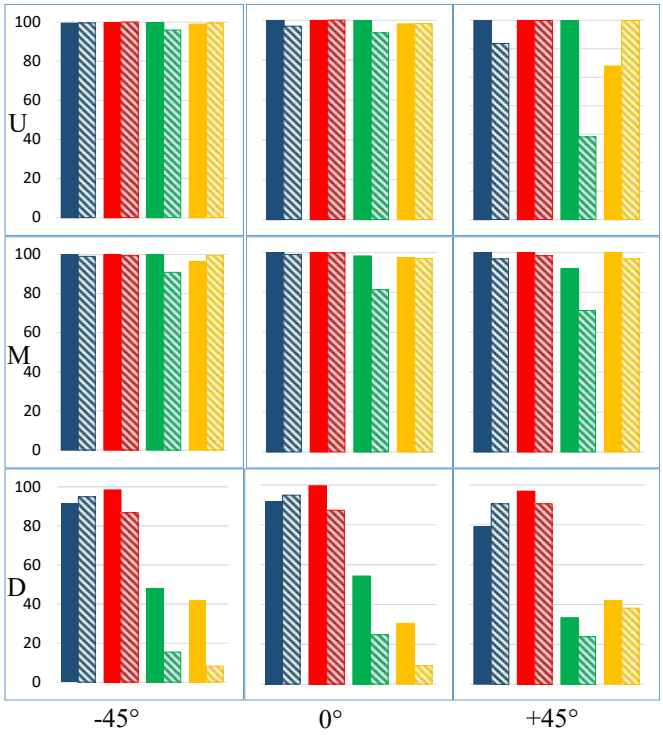

Fig. 4 GAR of the four SDKs (Sdk1: blue, Sdk2: red, Sdk3: green, Sdk4: yellow) for the *Lighting* testing set before (asolid bar) and after demorphing (dashed bar). The results are reported as a function of the elevation (D, M, U) and the azimuth (-45°, 0°, +45°) of the lighting source.

Face appearance is noticeably altered by these lighting variations as confirmed by the GAR variations in Fig. 4. In particular, the SDKs suffer when the face is illuminated from below (D), independently of the azimuth. This kind of unnatural lighting can make the detection of the main facial landmarks critical and compromise the whole recognition process. As to the effects of face demorphing, the results show that where the SDK is reliable and robust the demorphing process does not affect the performance significantly; on the other hand, in the cases where the performance of the SDKs are natively not good, face demorphing causes a further reduction of the accuracy. However, it is worth noting that, in the eMRTD scenario, extreme illuminations are quite unlikely because the gates are installed and setup to maximize the recognition performance.

Pose variations have been evaluated on the *Pose* testing set of the CAS-PEAL-R1 database. The images have been acquired by varying the pose of the subject in terms of both elevation (U: looking up, M: looking at the camera, D: looking down) and azimuth (in the range [-45°;+45°]). Higher values of azimuth are not relevant for this scenario, even considering that the user is interested in cooperating with the system to be quickly recognized. Since the CAS-PEAL-R1 databases does not contain testing images for the frontal pose (elevation "M" and azimuth +00), we used for this test the images of the *Aging* subset containing natural and frontal images acquired about six month later than the first acquisition.

ISBN 978-90-827970-1-5 © EURASIP 2018

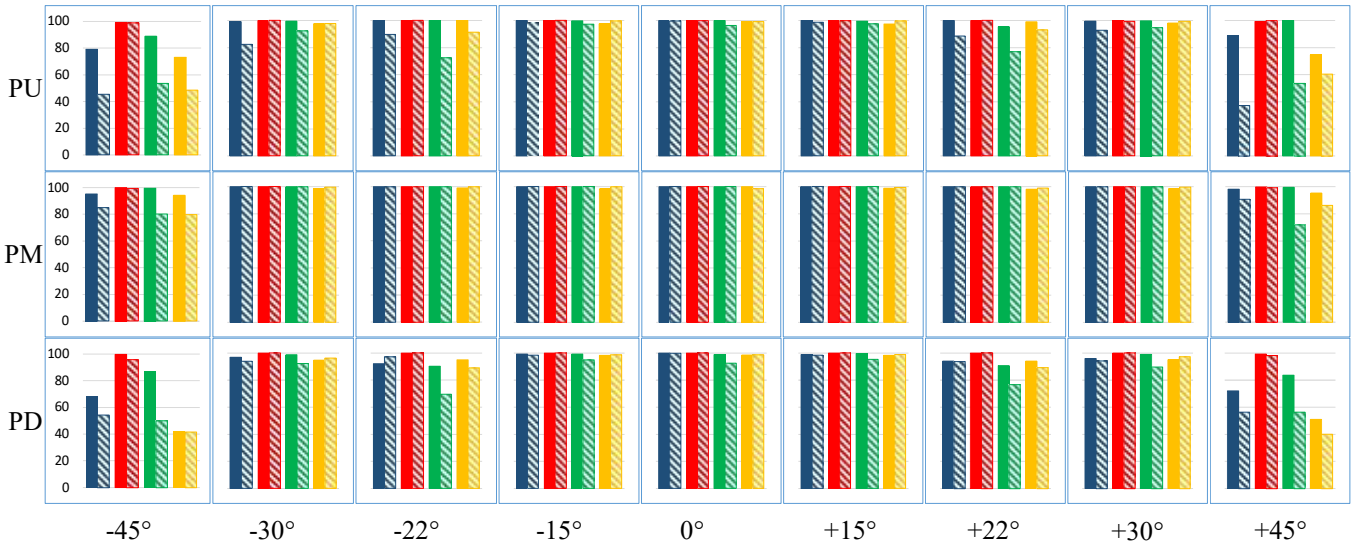

Fig. 5 GAR of the four SDKs (Sdk1: blue, Sdk2: red, Sdk3: green, Sdk4: yellow) for the *Pose* testing set before (solid bar) and after demorphing (dashed bar). The results are reported as a function of the elevation (PD, PM, PU) and the azimuth [-45°, +45°] of the lighting source.

The results obtained on a set of 20187 comparisons are summarized in Fig. 5 and Table II. For this specific variation, besides the GAR (see Fig. 5), the number of comparison failures (Table II) must be considered as well. In fact, unlike for other tests, in this case, the failure rate is quite high for some SDKs and this must be tracked.

TABLE II - COMPARISON FAILURE COUNT FOR THE FOUR SDKS (BEFORE DEMOPHING) AS A FUNCTION OF THE SUBJECT'S POSE.

	E	*Azimuth*								
		-45	**-30**	**-22**	**-15**	**0**	**+15**	**+22**	**+30**	**+45**
Sdk1	**U**	0	0	0	0	0	0	0	0	0
	M	0	0	0	0	0	0	0	0	0
	D	0	0	0	0	0	0	0	0	0
Sdk2	**U**	7	0	0	0	0	0	0	0	7
	M	0	0	0	0	0	0	0	0	1
	D	21	1	1	0	0	0	0	1	21
Sdk3	**U**	941	533	43	50	9	56	58	551	996
	M	767	141	0	0	0	2	3	209	826
	D	992	512	29	56	10	65	37	529	985
Sdk4	**U**	14	14	3	14	14	13	4	15	15
	M	13	13	0	5	1	9	0	13	11
	D	16	10	0	12	8	12	2	16	10

For limited head rotations (i.e., an azimuth in the range [-15°;+15°]) the failure rate is overall quite limited and the accuracy is not much affected for most of the SDKs; some of them work well even at a wider range of poses ([-30°;+30°]), even if the failure rate exhibits a growing trend. The most extreme poses corresponding to a profile face (head rotation of 45°) are difficult to address for most SDKs; in some cases a very high failure rate is observed (e.g., Sdk3), in others a noticeable performance drop is registered. Sdk2 is very robust to this variation as well reaching almost 100% accuracy with a very limited number of failures.

This variation represent a challenge for face demorphing which heavily relies on the alignment between corresponding points in the two images; in fact, such alignment is quite difficult to be performed in the presence of so different poses. This difficulty is particularly evident at the most extreme poses (head rotation of 45°). When the verification accuracy of the SDKs is lower (Sdk1, Sdk3 and Sdk4) the loss introduced by face demorphing is quite relevant; on the other hand for Sdk2, which is very robust also in the presence of pose variations, the demorphing process seems to not alter the performance. This means that the images produced by face demorphing are still recognizable and visually acceptable even in these extreme conditions, but probably some SDKs are sensitive to the artifacts generated as a consequence of face misalignment.

IV. CONCLUSIONS

The outcomes of this study are quite encouraging: face demorphing proved to be robust to a wide range of alterations and, in many cases, its application does not alter the SDKs accuracy. The most extreme lighting and pose variations represent a challenge for face recognition in general; the effect of face demorphing in these cases further compromise the accuracy. In a real scenario, however, such extreme variations are quite unlikely for two reasons: i) the image acquisition setup is designed to maximize the recognition accuracy and ii) the travelers are usually cooperative thus avoiding strange poses and expressions. The automatic detection of particular image conditions (e.g., excessive head rotation or uneven lighting) could be set up to avoid issuing false morphing warnings.

In general, we observe that the accuracy of face demorphing is strictly related to the intrinsic robustness of the face verification SDK; for instance Sdk2 is very robust to all the variations and the performance is not significantly altered by face demorphing.

While this work provides important empirical evidences to support the feasibility of demorphing for e-gate protection, further test sessions should be run in operational conditions,

ISBN 978-90-827970-1-5 © EURASIP 2018

and this is something we are going to setup in cooperation with partners in charge of airport security.

REFERENCES

[1] M. Ferrara, A. Franco, and D. Maltoni, "The Magic Passport," in *IEEE International Joint Conference on Biometrics (IJCB)*, Clearwater, Florida, USA, 2014, pp. 1-7.

[2] M. Ferrara, A. Franco, and D. Maltoni, "On the Effects of Image Alterations on Face Recognition Accuracy," in *Face Recognition Across the Electromagnetic Spectrum*. Switzerland: Springer International Publishing, 2016, pp. 195-222.

[3] M. Ferrara, A. Franco, and D. Maltoni, "Face Demorphing," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 1008-1017, April 2018.

[4] X. Zhang and Y. Gao, "Face recognition across pose: A review," *Pattern Recognition*, vol. 42, no. 11, pp. 2876-2896, November 2009.

[5] R. M. Makwana, "Illumination invariant face recognition: A survey of passive methods," *Procedia Computer Science*, vol. 2, pp. 101-110, 2010.

[6] K. Delac, M. Grgic, and S. Grgic, "Image compression effects in face recognition systems," in *Face Recognition*., 2007, ch. 5.

[7] NIST. (2018, March) Face Recognition Vendor Test (FRVT). [Online]. http://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt

[8] P. Grother, M. Ngan, and K. Hanaoka, "Ongoing Face Recognition - Part 1: Verification," NIST, Gaithersburg, MD, 2018.

[9] FRONTEX - R&D Unit, "Best Practice Technical Guidelines for Automated Border Control (ABC) Systems - v2.0," FRONTEX, Warsaw, Poland, ISBN: 978-92-95033-58-0, DOI: 10.2819/26969, August 2012.

[10] ISO/IEC 15444-1, Information technology - JPEG 2000 image coding system: Core coding system, 2016.

[11] ISO/IEC 19794-5, Information technology - Biometric data interchange formats - Part 5: Face image data, 2011.

[12] W. Gao et al., "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 38, no. 1, pp. 149-161, January 2008.

ISBN 978-90-827970-1-5