

Signature Recognition: Human Performance Analysis vs. Automatic System and Feature Extraction via Crowdsourcing

Derlin Morocho^{1,2(✉)}, Mariela Proaño¹, Darwin Alulema¹,
Aythami Morales², and Julian Fierrez²

¹ Universidad de las Fuerzas Armadas – ESPE, Sangolquí, Ecuador
{dmorocho, mcproano6, doalulema}@espe.edu.ec

² ATVS – Biometric Recognition Group,
Universidad Autónoma de Madrid, Madrid, Spain
{aythami.morales, julian.fierrez}@uam.es

Abstract. This paper presents discriminative features as a result of comparing the authenticity of signatures, between standardized responses from a group of people with no experience in signature recognition through a manual system based on crowdsourcing, as well as the performance of the human vs. an automatic system with two classifiers. For which an experimental protocol is implemented through interfaces programmed in HTML and published on the platform Amazon Mechanical Turk. This platform allows obtaining responses from 500 workers on the veracity of signatures shown to them. By normalizing the responses, several general features which serve for the extraction of discriminative features are obtained in signature recognition. The comparison analysis in terms of False Acceptance Rate and False Rejection Rate finds the presented features, which will serve as a future study of performance analysis in the implementation of automatic and semiautomatic signature recognition systems that will support financial, legal and security applications.

Keywords: Amazon Mechanical Turk · Workers · FAR · FRR · Crowdsourcing

1 Introduction

Since the beginning of humanity, it has been a need to identify and verify people by various means. One way to check the identity of people is the handwritten signature, which corresponds to a biometric feature. Biometric verification is a very important topic of research and is focused on identity verification applications. Biometric measurements are used for security purposes, since these are not easy to duplicate, and also cannot be stolen or forgotten, so they are safer [1]. Because of the warranty presented by written signatures, it has been seen today the need to build systems that perform biometric verification of signatures' strokes. There are automatic signature verification systems that require a lot of training information for proper performance. At present, many activities of daily life require a signature to identify an individual. This identification is performed by humans with no experience in forensic document analysis in

most applications such as a bank or a public notary, where time and decision-making depend on having agility and ability to observe minor variations.

To massively collect data in signature recognition by humans, crowdsourcing, is the tool that allows such activities, which is implemented through Amazon Mechanical Turk (MTurk). At present, crowdsourcing is a tool that uses a multitude of human beings to solve different types of problems, in the state of the art there are research projects that use this method, such as face recognition [2], gait recognition [3], and biometric security [4] for which this method is an alternative for signature recognition. Face recognition is used in applications such as mobile device authentication [5] and site security verification [6]. To run these applications, crowdsourcing helps determining human performance regarding algorithms and a COTS face comparator (Commercial-off-the-shelf) [2], concluding that humans have a better performance in terms of False Acceptance Rate (FAR) with a low value of False Rejection Rate (FRR) [2]. This is because human beings have the ability to use contextual information known as soft biometrics (ethnicity, gender, hair profile, etc.) for recognition of strokes and faces [3].

Human recognition capabilities were tested to identify people and scenarios in low quality surveillance videos. The recognition is based on human perception to tag subjects in various environments. Humans have proven to be objective, outstanding, reliable and robust to changes in viewing distance, as mentioned in [3].

MTurk is a platform used in several studies. Currently there are security systems based on voice recognition [7]. In [4] a method for identifying participants that execute mimics for different speakers of a target population is proposed.

The performance of people doing signature recognition is studied in [8, 14] and shows promising results in human performance as they are being able to distinguish a genuine signature from a forged one. Due to the positive results of these studies, it is proposed to use the tool of crowdsourcing implemented in MTurk to attack the problem of intravariability to make handwritten signature recognition with human intervention.

The article is distributed as follows: Sect. 1 shows the related works and importance of this paper, Sect. 2 shows state of de art of Automatic and manual recognition systems. The analysis and performance comparison are analyzed in Sect. 3 and finally the conclusions are discussed in Sect. 4.

2 State of the Art of Automatic and Manual Recognition Systems

2.1 Crowdsourcing via Amazon Mechanical Turk

Crowdsourcing is a participatory type of online activity in which an individual, institution, nonprofit organization or Company proposes to a set of people with different kinds of knowledge, diversity and number, through a free call, a task to be carried out voluntarily. The execution of the task, of varying complexity and modularity, and in which the crowd should participate by contributing with their work, money and knowledge, always involves a mutual benefit. The advantages of this tool

such as large numbers of participants, maximized performance, fast response rate and low costs of operation, have allowed researchers to have good results that have seen the human potential regarding to automated systems [2–4].

To implement crowdsourcing, it is required the use of MTurk [10]. Due to the features shown by MTurk, this platform has been used in face recognition [2], the obtained results show that humans have a superior performance compared to machines because they are capable of using contextual information of the image.

2.2 Manual Handwritten Signature Recognition System Based on Crowdsourcing

Manual signature recognition is performed by forensic document analyzers which have a method for studying handwriting features as mentioned in [13]. Manual signature recognition system is formed by the requester, HIT, MTurk platform, workers and the obtained results. Figure 1 shows how the manual signature recognition system is formed.

The requester is responsible for designing the HIT according to the experimental protocol to be evaluated. HITs are tasks that require some level of intelligence for its completion and are implemented using HTML. MTurk allows requesters to publish HITs which are solved by workers to obtain the results.

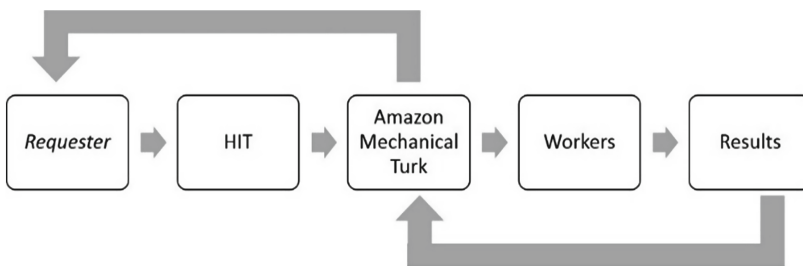


Fig. 1. System based on crowdsourcing

For manual recognition based on crowdsourcing, it is required to implement a GUI that contains instructions, training and test signatures and answer options. The interface shows four training signatures and one test signature, where the worker observes training signatures and qualify the test signature giving a similarity value between 1 and 10. Figure 3 shows the activities the workers should perform to evaluate the signatures (Fig. 2).

The aim of the protocol is to compare the performance of the human in a possible real scenario with more information. The protocol configuration parameters are shown in Table 1 (Fig. 4).

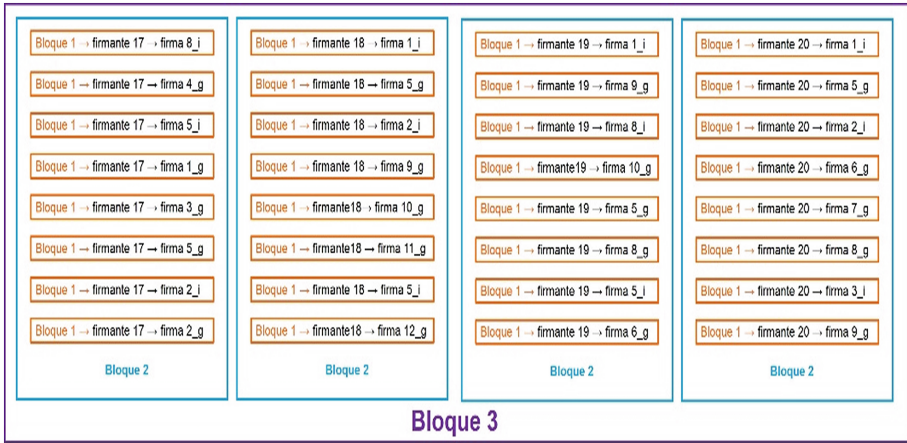


Fig. 2. Protocol interface with 8 test signatures for 4 signers

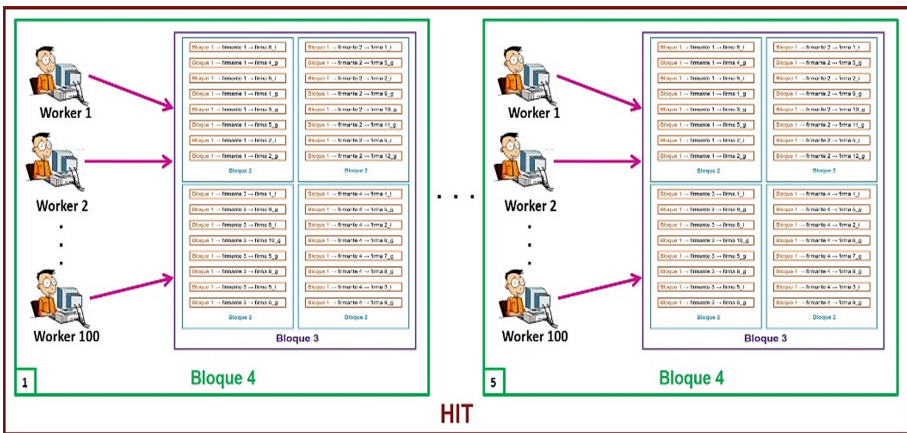


Fig. 3. HIT for 500 workers with 20 different signers

Table 1. Data configuration of the crowdsourcing system

Parameter	Configuration
Database	BiosecurID
Signers	20
Execution time	3 min
Number of workers	500
Number of activities	4
Number of analyzed signatures	240 divided into: <ul style="list-style-type: none"> - 80 genuine training signatures - 160 genuine and forged test signatures

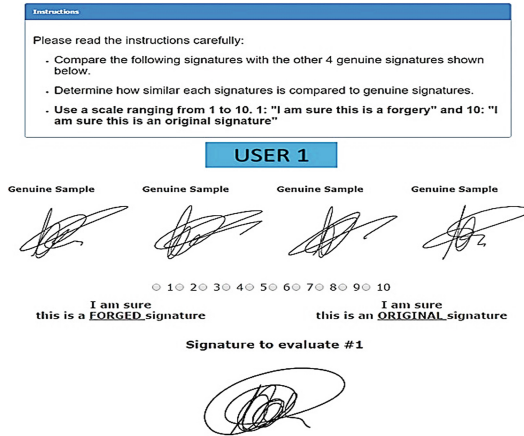


Fig. 4. Protocol interface with the first test signature

2.3 Automatic Recognition System

The recognition system used for the proposed comparison performance [11], uses global and local features. To classify the local features a Back Propagation Neural Network (BPNN) was used while global features use the probabilistic model. In order to use the recognition benefits from both classifiers, an “AND” combination of the classifiers is done to make the final decision [11] (Table 2).

Table 2. Configuration data of the automatic system

Parameter	Configuration
Database	SVC2004
Signers	40
Number of analyzed signatures	600 divided into: - 200 genuine training signatures - 400 genuine and forged test signatures

The global features’ classifier uses the probabilistic model for its development. First, feature values of the training signatures are obtained, then its mean and variance is calculated and thus a threshold value is defined. This threshold is the lowest value of the training signature’ features, which is compared with the probability score (PS). If the PS value is greater than or equal to the threshold value, the user is considered genuine, otherwise is rejected as a forgery [11].

The local features’ classifier corresponds to a BPNN with a 3-layer MLP and in which the output of the neural network is evaluated based on the desired output. If the response is not satisfactory, the connections (weights) between layers are modified and the process is repeated until the error is low [12]. In this case the threshold value (desired output) is 0.5. If the values obtained from the output of the BPNN is greater than or equal to 0.5 the user is considered genuine, otherwise it is rejected as a forgery [11].

In [11] the logical operation “AND” to combine the two classifiers is explained and its benefits are shown. Thus, a user is considered genuine if it is recognized as genuine in both classifiers, otherwise he/she is considered as a forger.

For performance comparison, the values of FAR and FRR from the combination of both classifiers and the probabilistic classifier are used [11].

3 Performance Comparison and Results’ Analysis

Performance comparison is accomplished by using data obtained from online recognition system based on two classifiers and manual system based on crowdsourcing.

The manual system based on crowdsourcing uses the BiosecurID database and the comparison for 20 signers is done. The automatic system uses de SVC2004 database for 40 signers.

The FRR and FAR values used correspond to those obtained from the combined responses of workers and the decision obtained from the combination of the two classifiers. The combination of responses of the manual system based on crowdsourcing is done by obtaining FAR and FRR mean values of several workers taken randomly, as shown in (1). Furthermore, 10 repetitions of this process are made, afterwards the mean value of the obtained results in repetitions is calculated as shown in (2).

$$cr = \frac{1}{n} \left(\sum_{i=1}^n w_i \right) \quad (1)$$

$i = 1, 2, \dots, n$ Random number of answers to combine
 w Worker’s answer in terms of FAR or FRR
 cr Mean value of n combined answers

$$er = \frac{1}{m} \left(\sum_{j=1}^m cr_j \right) \quad (2)$$

$j = 1, 2, \dots, m$ Number of repetitions
 cr Mean value of n combined answers
 er Evolution of the m repetitions of combined answers

In this case the number of combined responses ranges from 1 to 500 in the manual system. Table 3 shows the values of FRR and FAR of the variations of the automatic and manual systems.

Table 3. FFR and FAR response performance in handwritten signature recognition

Method	FRR (%)	FAR (%)
1 Automatic probabilistic (40 signers)	27	9
2 Automatic combined (40 signers)	3	5
3 Manual crowdsourcing (20 signers)	32	38
4 Manual Crowdsourcing Combined (20 signers)	7	24

Table 3 shows that Manual Crowdsourcing Combined System compared to Automatic probabilistic system in terms of FRR is better, thus the Manual Crowdsourcing Combined system rejects less genuine signatures, demonstrating the ability of humans in signature verification, but in terms of FAR, the Manual Crowdsourcing Combined system, admits a greater number of impostors firms, implying that the human is more permissible to accept impostors signatures as genuine.

When analyzing the FRR of Automatic Probabilistic, it can be observed that it is slightly lower than the value of Manual Crowdsourcing system, this shows that both, human and machine, when working without combinations may wrongly reject genuine signatures. However, the value of FAR is different because automatic probabilistic system shows superior performance compared to Manual Crowdsourcing system. This shows that the automatic probabilistic recognition accepts less forged signatures as genuine.

The Manual Crowdsourcing system has a FRR value (32 %) and FAR value (38 %) higher than the automatic combination (3 % FRR and 5 % FAR) because it uses for signature recognition, people that are not experts in forensic document examination. Figures 5 and 6 show the evolution of FAR and FRR respectively when workers' responses are combined in signature recognition.

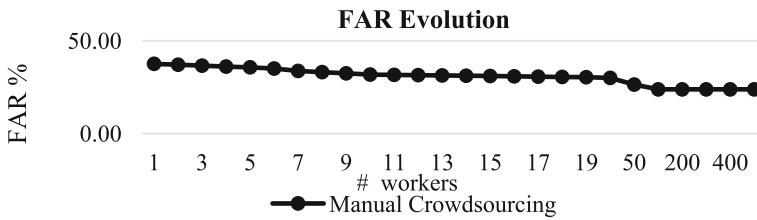


Fig. 5. Performance evolution FAR

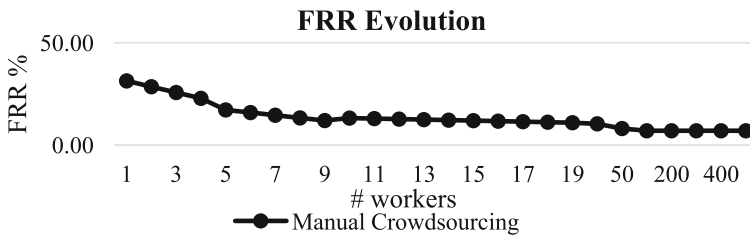


Fig. 6. Performance evolution FRR

Figures 5 and 6, show that FAR = 7 % and FRR = 24 % respectively, stabilize when the answers of 100 workers are combined in handwritten signature recognition. Manual Crowdsourcing System provides information of discriminative features for the identification of a handwritten signature. Figure 7 shows the signature features and the percentage of workers who used each one of them.

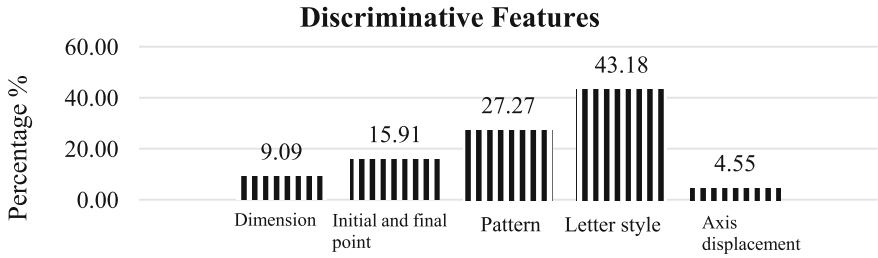


Fig. 7. Features extraction defined by workers

In Fig. 7, it can be seen that **Letter style**, is the characteristic chosen by the workers to determine whether the signature is genuine or forged. Furthermore, the **Axis displacement** characteristic is the less applied in signature recognition.

From the discriminative features of Fig. 7, a subdivision of them is extracted and is shown in Figs. 8, 9 and 10. It is noted that the traits present in **Letter style**, are the most used by workers in handwritten signature recognition, this happens because humans are able to detect minute variations in features due to their perception [3].

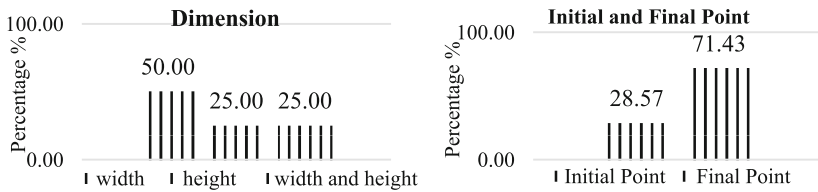


Fig. 8. Features: dimension and initial and final point

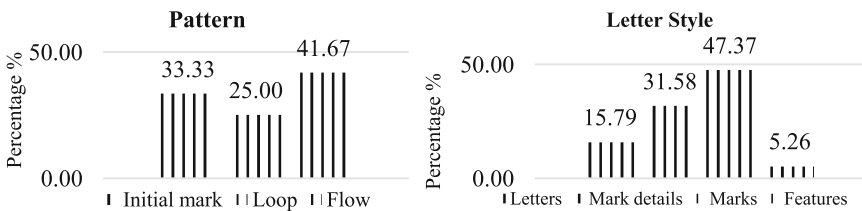


Fig. 9. Features: pattern and letter style

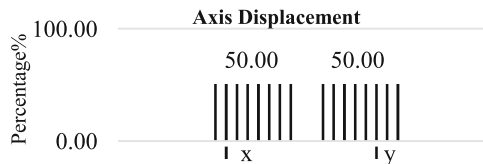


Fig. 10. Features: axis displacement

The discriminative features obtained through crowdsourcing and the results of this classification, generate a set of 7 features discriminating for handwritten signature recognition. Tables 4 and 5 show and briefly describe the 7 discriminative features.

Table 4. Discriminative features 1

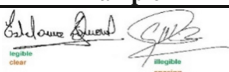




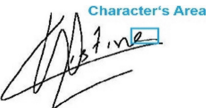

Characteristic	Definition	Example
Legibility	It refers to the existence of defined literal characters	
Total area occupied by the signature	The most relevant signatures' points are taken (4 strategic points containing signature) to calculate the area of the signature	
Perimeter	It is proposed to measure the contour of the signature that best represents it.	
Pressure of initial and final points	It is propose the identification of signatures' initial and final points and the measurement of pressure of these points.	

Table 5. Discriminative features 2

Characteristic	Definition	Example
Connection between characters	The signature is observe to determine whether it is legible or illegible. If the signature is illegible then it is identified the existence of relevant connections between characters. If there are connections they should be classified in rounded, sharp or mixed.	
Character Area	Measuring the area of a character through a box is proposed. The box should contain inside the character which desired area is being measured.	
Proportionality	The signature strokes are observed and height levels representative points are defined, with the aim of establishing a dependent dimensionless ratio in the signature	

4 Conclusions

According to research and analysis of responses from the workers, these alone are not competent and able to recognize signatures. Static information displayed to the workers for signature recognition is not sufficient and appropriate to improve performance. However, by combining the responses of each worker, performance has improved by 40 % in FAR and FRR by 80 %.

The criteria from workers' responses allow extracting discriminative features of a signature such as: dimension, initial and final points, pattern, letter style and displacement, permitting the release of new features that will help automatic signature recognition systems and in the future be support for generating semiautomatic signature recognition systems.

Performance results of manual system based on crowdsourcing envision a bright future because, by combining the answers it may give a better performance to automatic and semiautomatic signature recognition systems. It is noted that the results are comparable with an automatic classifier based on a probabilistic model, therefore, merging the benefits of a manual and an automatic system could lead to the development of a high performance automatic system with human oversight. In addition, the combination of workers' responses in HITs, provide future research, which would be the analysis of discriminant features that provide better performance.

This paper is the beginning of a new research that includes HIT generation in signature recognition by improving or establishing new parameters such as: number of signers, number or workers, comparison levels, focusing on new scenarios of human assistance applied in signature recognition through crowdsourcing.

References

1. Toscano Medina, L.: Reconocimiento Dinámico y Estático de Trazos (2009)
2. Best-Rowden, L., Han, H., Otto, C., Klare, B., Jain, A.: Unconstrained face recognition: identifying a person of interest from a media collection. *IEEE Trans. Inform. Forensic Secur.* **9**(12), 2144–2157 (2014)
3. Martinho-Corbishley, D., Nixon, M., Carter, J.: Soft biometric recognition from comparative crowdsourced annotations. In: 6th International Conference on Imaging for Crime Prevention and Detection (ICDP 2015). IET, pp. 1–6 (2015)
4. Panjwani, S., Prakash, A.: Crowdsourcing attacks on biometric systems. In: Symposium On Usable Privacy and Security (SOUPS 2014), 2014, pp. 257–269 (2014)
5. Sandoval, A. López, E., Martínez, C., Cruz Rivas, L.: Sistema de Autenticación Facial mediante la Implementación del algoritmo PCA modificado en Sistemas embebidos con arquitectura ARM. *La Mecatrónica en México*, **4**, pp. 53–64, 2015
6. Andrade, C.N.: Autenticación por reconocimiento facial para aplicaciones web, utilizando software libre (2015)
7. Pérez Badillo, O., Poceros Martínez, F., Villalobos Ponce, J.: Sistema de seguridad por reconocimiento de voz (2013)
8. Morocho, D., Morales, A., Fierrez, J., Tolosana, R.: Signature recognition: establishing human baseline performance via crowdsourcing. In: Proceedings of 4th International Workshop on Biometrics and Forensics (IWBF). Limassol, Cyprus (2016)
9. Estelles-Arolas, E., Gonzalez-Ladron-de-Guevara, F.: Towards an integrated crowdsourcing definition. *J. Inf. Sci.* **38**(2), 189–200 (2012)
10. Amazon Mechanical Turk, [En línea]. Available: <http://www.mturk.com>. Último acceso: 3 Noviembre 2015
11. Alhaddad, M.: Multiple classifiers to verify the online signature. *World Comput. Sci. Inf. Technol. J. (WCSIT)* **2**, 46–50 (2012)

12. Cilimkovic, M.: Neural Networks and Back Propagation Algorithm. Institute of Technology Blanchardstown (2008)
13. Harrison, D., Burkes, T., Seiger, D.: Handwriting examination: meeting the challenges of science and the law. *Forensic Sci. Commun.* **11**(4) (2009). <https://www.fbi.gov/hq/lab/fsc>
14. Morocho, D., Morales, A., Fierrez, J., Vera-Rodriguez, R.: Towards human-assisted signature recognition: improving biometric systems through attribute-based recognition. In: *IEEE International Conference on Identity, Security and Behavior Analysis (ISBA 2016)*. Sendai, Japan (2016)