das-FaceBond

Biometric Identification Software for Identity Fraud Prevention

Revisión	Fecha	Descripción	Redactado	Revisado	Aprobado
1	12/06/2018		SGP	EMR	EAL

1. Introduction

In the facial biometrics field,two situations are typically handled:

- **Identification** 1:N or M : N, in order to verify the identity of the new customers against the current customer's database.
- **Clustering**:in order to detect multiple identity fraud cases in the current customer's database.

das-FaceBond, incorporates the latest facial biometric technologies that make possible to work with both faces galleries and ID documents.

das-Face, the facial biometric engine, has been compared to the best biometric engines in the world under the conditions set by the LFW database (13.223 pictures of 5.749 different persons). The performance results of Veridas´ facial biometrics engine are detailed in the chart below:

Face Biometry System	Accuracy (%)
Baidu ensemble model [1]	99.77
Baidu single model [1]	99.68
das-Face	**99.67 ± 0.30**
FaceNet (Google) [2]	99.63 ± 0.09
Human (overestimated)	97.53
DeepFace (Facebook) [3]	97.35 ± 0.25

[1] J. Liu, Y. Deng, and C. Huang. "Targeting ultimate accuracy: Face recognition via deep embedding". arXiv:1506.07310, 2015.
[2] F. Schroff, D. Kalenichenko, and J. Philbin. "Facenet: A unified embedding for face recognition and clustering". CVPR, 2015.
[3] Taigman, Y., Yang, M., Ranzato, M. & Wolf, L. "Deepface: closing the gap to human-level performance in face verification". Proc. Conference on Computer Vision and Pattern Recognition 1701–1708 (2014).

2. Das-FaceBond

2.1. Identification System

Das-Face is the facial biometrics engine proprietary of Veridas that has been developed using Deep Learning and AI technologies. It is defined as Identification the process of searching a person or a gallery of persons within a database of identities.

The different types of identification are detailed below:

- **Identification (1:N):** given a certain person to be found within an image gallery or a database with multiple persons (N), the system provides the most similar image (1) to the target person within the database.

- **Identification (M:N):** given a set of images of different persons (M) and a gallery or database with multiple images of different persons (N), the system provides all the images from the database (N) that correspond in terms of similarity to the images of the input gallery (M).

- **Clustering:** given a database containing multiple images of different persons (N), the system generates a cluster of images for each single identity found within the database (N).

The identification system would enhance the following tasks:

- Create and delete as many galleries as desired:
  - Documents galleries
  - Image galleries
- Add and delete pictures of a certain gallery. One by one or all together (*batch*).
- Visualize the different identities of each gallery. For each identity, its own picture can be visualized including a set of parameters linked to the identity (email, name, ID number, location, etc).
- Possibility of searching a particular person in a gallery through its meta_key. Thanks to this, a person can be found by introducing its name (for example to delete him from the gallery).
- Check the historic data of the identification processes.
- By Clicking on each identification process, the system will give access to a detailed screen in which the best match and the following "N" matches can be checked, being "N" adjustable. In other words, the best (10,100...) matches can be seen.
- Possibility of filtering the identification processes by meta_key or by date (including hour and minute in the filtering criteria).

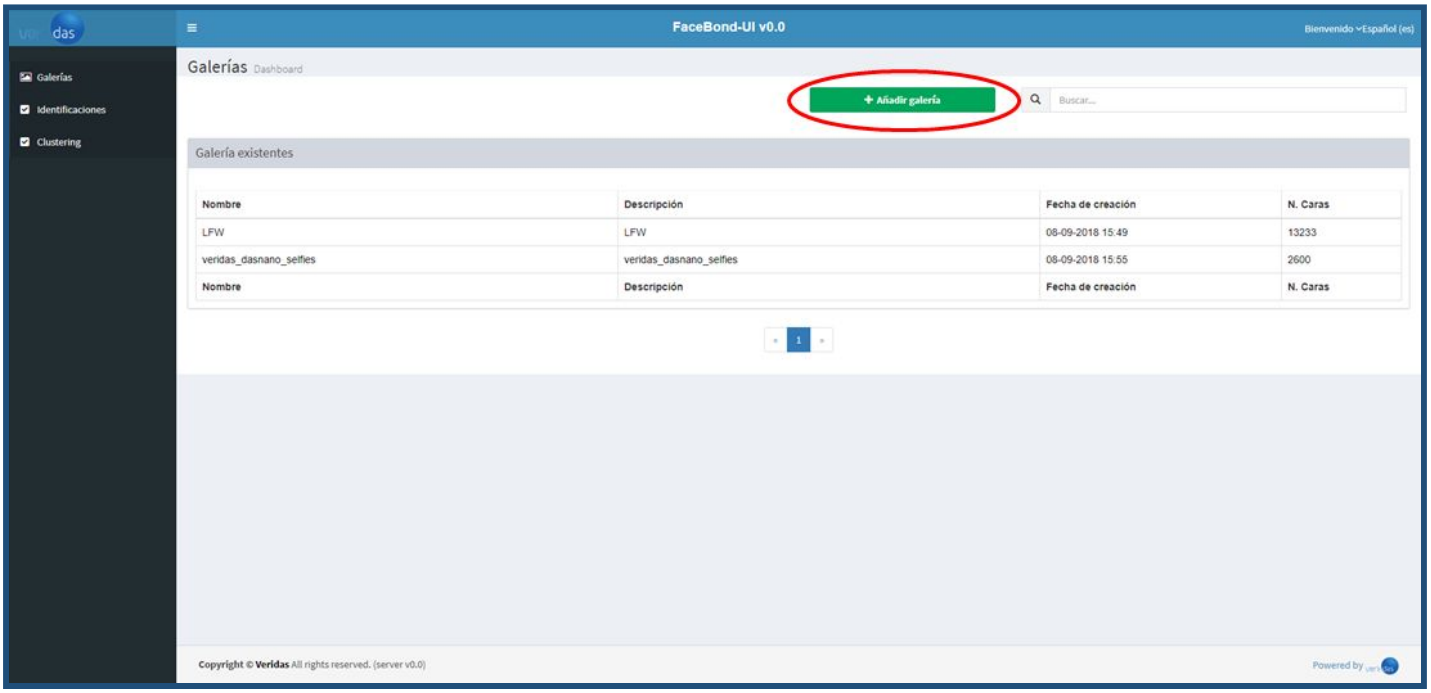
The input images can be loaded from different sources:

- Any portrait of a person, independently of the source.
- Any ID document containing a person' picture
- Printed images
- Image present in the NFC (if existing) from different documents: DNI 3.0, passports with ICAO normative (e-passports).
- Photograms of a video Stream.
- Veridas Capture SDK's: mobile and HTML
- **Batch of files (.zip or .tar)**

3. User Interface

The different screens of the user interface are shown in this chapter:

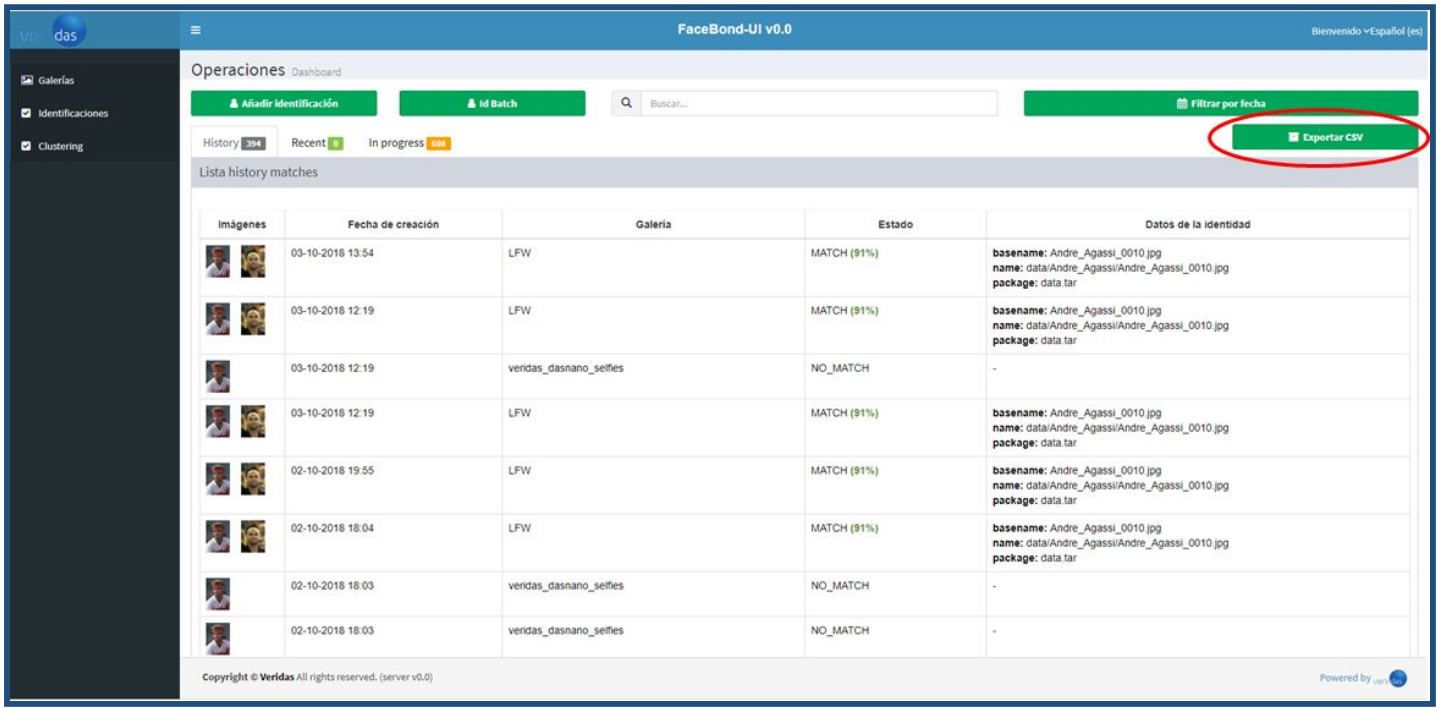
3.1. Main interface



This figure shows the main interface of the biometric identification software in which different sections such as gallery, identification or clustering can be accessed.

Adding a new gallery is available, clicking on the button highlighted in red.

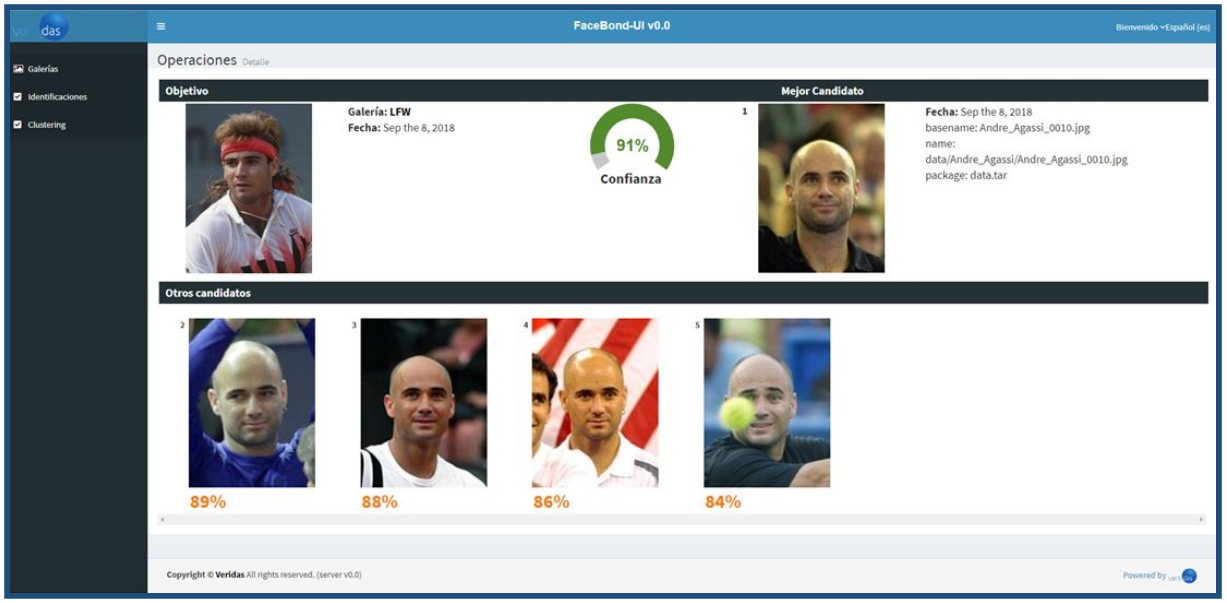
3.2. Identification interface



This figure shows the identification interface in which both single person images or batch of images can be uploaded.

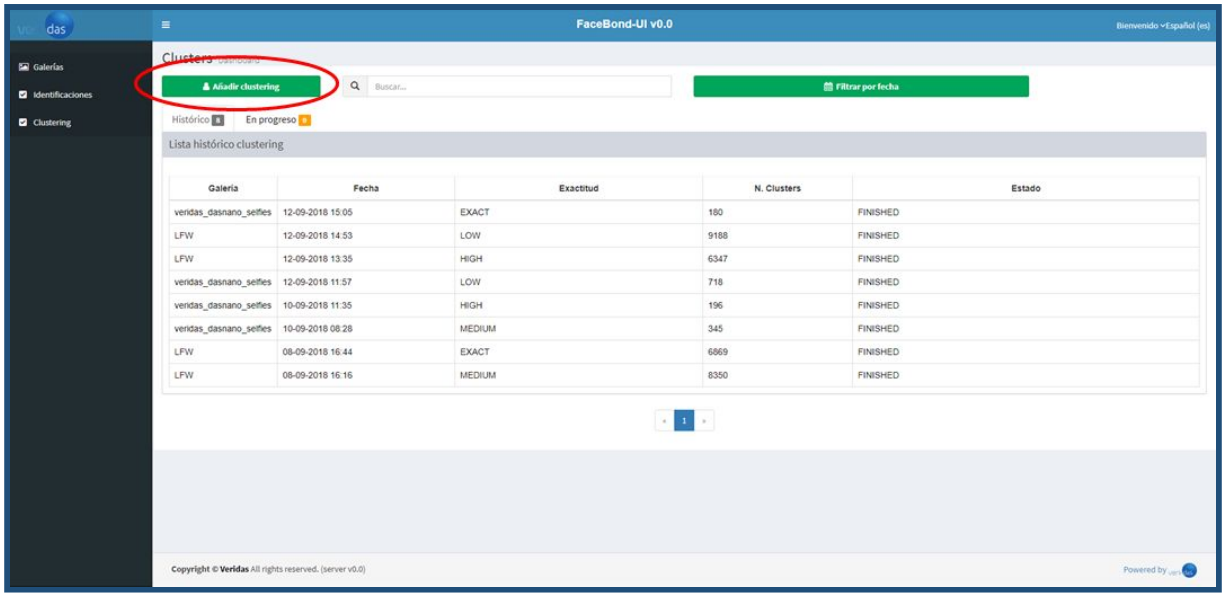
Exporting the data is available clicking on the button highlighted in red.

3.2.1 Identification results interface



This figure shows the identification results interface where the matches of all the identification processes carried out can be found.

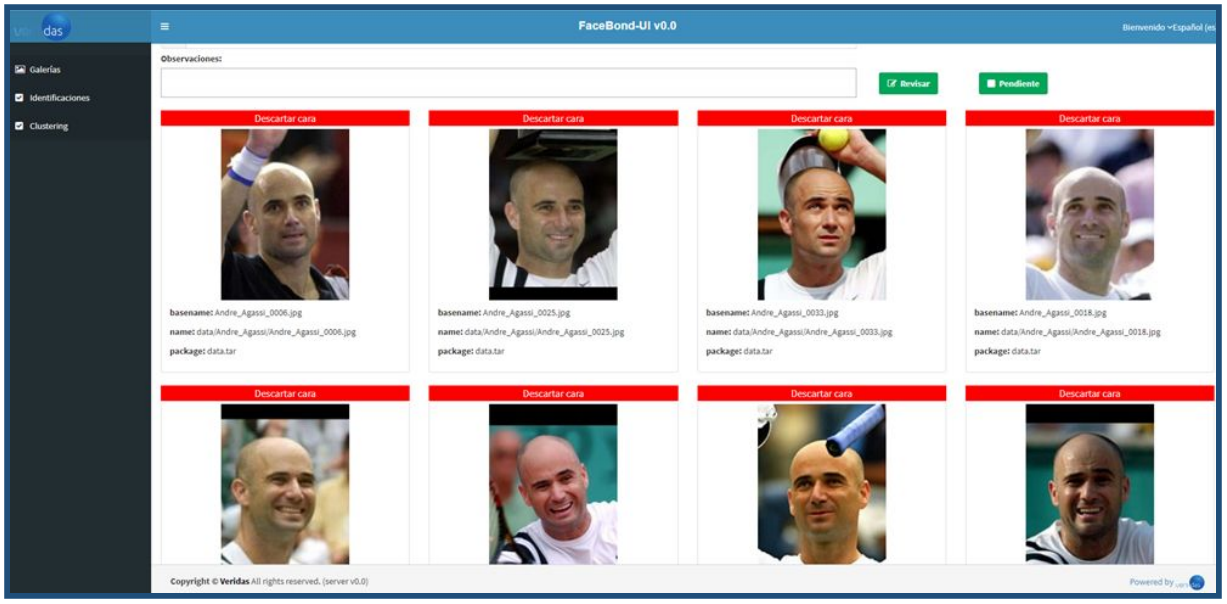
3.3. Clustering interface



This figure shows the clustering interface in which a new cluster can be added or searched.

Creating a new cluster is available clicking in the button highlighted in red.

3.3.1. Clustering results interface



This figure shows the clustering results interface in which all the obtained matches are shown.

This cluster can be corrected manually. If any inconclusive image has been grouped, it can be deleted pressing the red box on top of each picture.

4. Technical Specifications

4.1. System Components

The service das-FaceBond requires a network of different services for running all the features of the system:

- das-Face docker version >= 1.13.0.
  - It is recommended to run it with at least 4 workers with a GPU available.
- Redis docker version >= 4.0.
  - No special configuration is required.
- PostgreSQL docker version >= 9.6.1.
  - It is recommended to use an SSD disk for the database.
- das-FaceBond docker version 1.13.0.
  - It may run with at least 500MB of memory per worker.
- Celery container running inside a different das-FaceBond docker 1.13.0 image.
  - Depending on your parallel requirements, more than one celery instances may be required.
  - The should be executed with at least 1GB of memory per celery worker.

At the end of this document you may find a docker-compose YAML description which puts all these containers up on a single server. Use it as an example for your own deployments.

4.2. Hardware Requirements

The minimum requirements for production purposes are:

- Machine with at least 4GB de RAM, and 4 CPU cores at 2GHz for das-FaceBond server.
- Machine with at least 8Gb de RAM with 16 CPU cores at 2GHz for Celery workers.
- SSD disk for database.

4.3. Accuracy and Performance

This system has been evaluated on a Intel(R) Core(TM) i9-7900X CPU @ 3.30GHz, with 64GB of RAM and the database in a SSD disk.

On identification, the system shows following accuracy and performance metrics:

- Accuracy is about 90% with MegaFace[1] dataset, using a gallery with 10,000 distractors.
- The system is able to search over 200,000 identification candidates per second.[2]

Regarding clustering operations, the system shows following accuracy and performance metrics:

- The clustering accuracy is about 97% with setting=EXACT and a minConfidence=0.95 on LFW[3] benchmark dataset.[4] LFW dataset contains more than 13,000 images.

[1] http://megaface.cs.washington.edu/

[2] This speed is without considering the time to generate the embedding vector for the probe face image.

[3] http://vis-www.cs.umass.edu/lfw/

[4] Consider this minConfidence just orientative, it is task dependent, and in some cases it will be required to use a higher value (0.98 or even 0.99).

- Clustering performance is shown in the following table, depending on the gallery size and the indicated accuracy level.

	Accuracy Level	
Gallery size	***EXACT***	***HIGH***
≅ 100	9 seconds	-
≅ 1,000	90 seconds	14 seconds
≅ 10,000	34 minutes	80 seconds
≅ 100,000	2 days	15 minutes

5. License

The following clauses set the terms, rights, restrictions and obligations on using this Software, created and owned by VERIDAS DIGITAL AUTHENTICATION SOLUTIONS, S.L. (the Licensor), without prejudice to the provisions laid down in the contracts subscribed by the Licensor and your entity (the Licensee), which shall prevail over this file.

5.1. License Grant

Licensor hereby grants to the Licensee a non-exclusive, non-assignable and non-transferable, indivisible, without the rights to create derivative works license to use this Software for the specific purpose specified between the parties, subject to the terms and conditions contained herein and other legal restrictions set forth in third party software used while running the Software.

The Software has different components that can be used for several applications. However, the License is granted over the Software licensed components as a whole, and no separated use is permitted other than the specific purposes agreed by the Licensor and the Licensee.

The Software is comprised of proprietary code. However, the Software may include certain third party components with separate legal notices or governed by other agreements. Even if such components are governed by other agreements, the disclaimers and the limitations on and exclusions of damages below also apply.

All intellectual and industrial property rights over and in respect of the Software are owned by the Licensor. The Licensee does not acquire any rights of ownership in the Software.

The use of reverse engineering practices is strictly forbidden.

5.2. Limitation of liability and indemnity

1. To the extent permitted under the law, the Software is provided under an "AS IS" basis. The Licensee acknowledges and agrees that neither the Licensor nor its board members, employees or agents, will be liable for any loss or damage arising out of or resulting from Licensor's provision of the Software under this License, or any use of the Software by the Licensee or its employees; and Licensee hereby releases Licensor to the fullest extent from any such liability, loss, damage or claim.
   Regarding the processing of personal data with the Software, Licensee acknowledges and agrees that Licensor is no liable for the data collected by the use of the Software within the scope of the Licensee's activities.
2. The Licensee must indemnify, defend and hold harmless the Licensor, its board members, employees and agents from and against any and all claims (including third party claims), demands, actions, suits, expenses (including attorney's fees) and damages (including indirect or consequential loss) resulting in any way from:
   - Licensee's and Licensee's employee's use or reliance on the Software;
   - any breach of the terms of the License by the Licensee or its employees;
   - any other act of Licensee that can be considered negligent.

3. The facial verification functionality of this Software has been tested against the LFW database, with results up to a 99.6% precision (FNR=0.5% at FPR≤0.5%). This level of reliability is subject to image capture conditions during the process, and therefore the Licensor shall not be responsible of any use the Licensee may make of the Software in different environments than those recommended by the Licensor.