Hybrid face recognition solution for security

Y Donon[1]

[1]Samara National Research University, Moskovskoe Shosse 34, Samara, Russia, 443086

Abstract. This article introduces a design that aims threw the combination of open source and closed source technologies, to make a, simple to implement,low-cost and high-performing face recognition solution. The solution provides identification, emotions and facial features recognition as well asdangerous objects spotting.This article exposes the concept of the solution, explains its importance on the market and provides details of a proof of concept prototype.

1. Introduction

The market of face and image recognition technologies is booming and forecasted a brilliant future. Although it is seen more and more in specialised magazine or promoted by giants of information technologies, many smaller actors are left behind as they perceive the technology as inaccessible or too expensive.

Numerous researchesaboutthose systems have been made in the recent times and during those years of research, computers science has evolved beyond measure.But what really have changed since a few years, are the cameras. What makes this ground of research more prolific than ever todays is that we all havephones in our pockets which sensors have an average of 14 megapixels, that we can buy full HD webcams for less than a hundred dollars. 15 years ago, a digital camera's resolution would be a fifth of what a webcam hasnow and be ten times its price. [1]

Although face recognitions attempts have been around for more than 50 years now, it still appears as a new technology to most of people. If we had indeed technologies able to perform those tasks back in the sixties, pictures would have to be taken according to very precise specifications. Attempts were multiplied; it became a trend in the nineties, some artefacts from that time, such as the ORL Database if Faces from Cambridge are even still in use today. In the beginning of the two-thousands, an international contest has even been thrown on the subject of face recognition. [2] Yet with all of that, it is only now and in the upcoming years that we really can and will perceive ground-breaking advance in those technologies. [3]

Nowadays, we have the tools, we have the necessary sensors for an efficient recognition and new actors on this market are emerging every day.Those solutions represent a trend on the security market of course; it allows to recognise not only people, but also specific objects and track them if necessary. The industry alsostarts to use emotion recognition systems to understand better their customer. [4] In this paper, I will introduce a solution to exploit this new market and make it accessible to everyone threw a low-cost, high performing face recognition solution for security. A design that is easy to deploy without high computing capabilities. The goal of this paper is for everyone to understand the stakes of this market, how accessible it is now and how it can be used in our everyday life.

2. Market and projections

2.1. Hybrid

As the market is still emerging but have been around for a long time, both open-source and closed source solutions exists. Closed source solutions are efficient to spot faces,can differentiate them, making an authentication possible. Those solutions, however, falls short when it comes to analyse a picture's details, such as emotions, facial details or objects. Closed source image recognition providers, however, are usually specialized and therefore extremely good when it comes to identifying those details. [4]

The design presented in this paper tries to take profit of this reality. Combining open source technologies and closed source ones, taking to both worlds what they are good at, allows making a first analysis on a local computer, even one having low computation capabilities and, over the internet, using solutions provided by the majors of image recognition, to analyse pictures in-depth,beyond the capacities of open source solutions.

2.2. Projection

As mentioned, the face recognition market is still emerging. It is expected to be worth between 7.5 and 10 billion dollars by 2022, 2 to 3 time more than it was worth in 2016. The year before that, the main client of those systems was US Homeland Security.By now the use of such solution for security has already spread in several countries and is used by such actors as the British police. Since its beginning this technology has been viewed as a major asset in security systems. [4]

Open source solutions are forecasted to improve their algorithms in 2D and thermal face recognition, while it is believed that online services will keep the specialized market (complex emotions, facial features details, 3D modelling, etc...) , although open source alternatives exists and will also improve, but not with the same precision rate. [5] The main uses between 2017 and 2022 are forecasted to be emotion recognition, tracking and monitoring, access control and law enforcement. [4] Therefore the design suggested in this paper fits the needs of the market to have an affordable solution, using the full capacities offered by the different actors of face recognitions solutions. It also is appropriate as thiscurrent is forecasted to be stable over at least the upcoming four years.

Making profitable for SMEs (Small and medium-sized enterprises), which are 98% of economic environment, a multi-billions digital economy market threw the design presented is a breakthrough for face recognition as it makes it an accessible tool.

3. Results

3.1. Functioning

In this design, if the picture is of sufficient quality for an optimal analysis[6], the system queries first a micro database of a handful of the most recent faces, loaded on the computer's RAM (1). This reduces the load on the disk's database and accelerates the program, as between two frames, it is usually the same faces that show up.

If the face hasn't been recognized on the first database, a query is sent to the second one, which can store up to a thousand of faces, depending of the capacities of the computer (2). This database is typically conceived to store the faces of all the employees of a company and manage access controls.

If no match is found in the second database (the confidency of the comparison between the shown face and the existing ones is too low), the system querries online services, that can analyse the picture, confirm that the idividual in unknown via an online database (3), and differenciate its emotions, facial features as well as alaysing its environement, detecting immediate threats, such as weapons.

Finnaly, the result of this detection will be added to the RAM-loaded database to avoid detecting and alaysing again the same face (4). Each query to an online service having of course a cost.

3.2. Results

The performance reached by the test program fitted all of our expectations, if sometimes the description suffers small imprecisions it offers a real-time identification on video with 5 frames per

second, spotting simultaneously several object[7], more than enough for a security camera, giving even an impression of relative fluidity in the capture. With identified facial features such as hair colour or emotions, a very precise recognition differentiating identical twins without any hesitation and beingable to detect some specific object such as weapons, we can say that, on a technical point of view, the performance test of the design is a complete success.

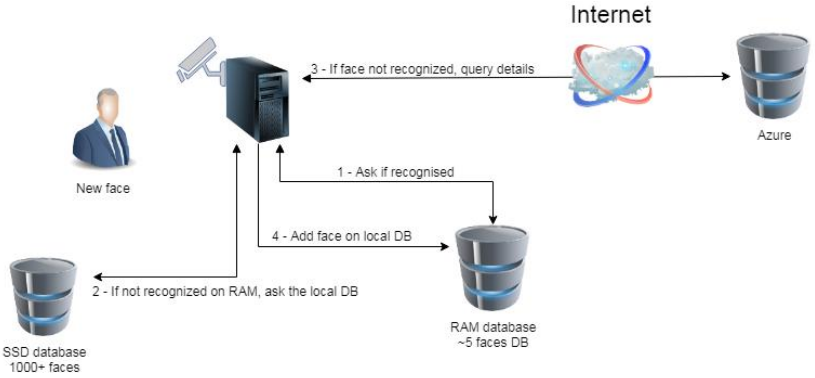

Figure 1. Design's business process.

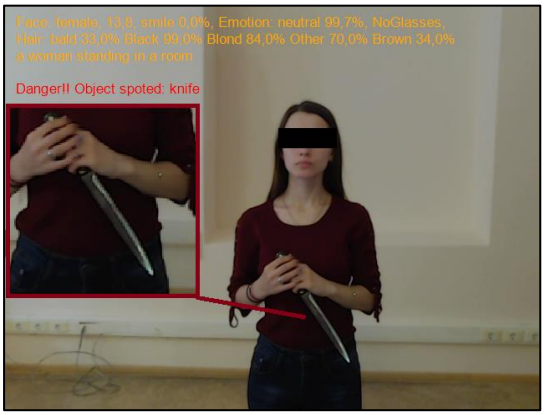

Figure 2. This image illustrates the analysis of a person caught on camera. Some of the information, such as the approximation of the age are not correct, however, the capture allows a clear identification (the program select the frame having the "best quality of face").

Figure 3. Assets that although the quality on an image might be poor, the program is able to extrapolate correct information.

Table 1. Features analysis for figure 2 and 3.

Feature	Values figure 2	Precision (appreciation if data unavailable) figure 2	Values figure 3	Precision (appreciation if data unavailable) figure 3
Facial	13.8	72%	20	97%
Smile	0.0%	Correct	0.0%	Correct
Emotion	Neutral 99.7%	Correct	Neutral 98.7%	Correct
Glasses	No glasses	Correct	Reading glasses	Correct
Hair	Bald 33%	15%	Bald 33%	15%
Hair	Black 99%	85%	Black 100%	Error
Hair	Blond 84%	10%	Blond 53%	Correct
Hair	-	-	Brown 42%	Correct
Hair	Other 70%	-	Other 38%	-
Description	A woman standing in a room	Correct	A woman in a blue shirt	Correct
Object	Knife	Correct	gun	Correct

Although some obvious progress are to be made on the hair colour detection, the features calculated are generally close to reality and most importantly allowsa human identification of a person, even without a the subsequent picture.

3.3. Technical specifications

A software has been developed as a design proof of concept. It has been developed in C#, using an OpenCV wrapper for this platform, OpenCvSharp, Fisherfaces recognition algorithm and Microsoft's Face and Vision API.

The use of Fisherfaces has been motivated over other methods for its search of discrimination criteria, which is more reliable to exclude possible faces match, enhancing the security offered by the solution. We widely favour a false negative, which leads to a control on the server that the person truly isn't identified in our database, than a false positive, which would allow an intruder to get through the system. [8-11]

The use of the Microsoft cognitive systems has been decided as it fitted the technical needs of the environment, offered a good transparency and as they send back details from the analysis of the image such as face coordinates, allowing further extrapolation. The other considered providers whichbilling systems were adapted to this design were Google Cloud Platform and IBM Watson.

The goal being to make the market as accessible as possible, it was important to reduce every source of costs. The system has been tested on several Microsoft Windows platforms (Win 7 and superior versions), it function and manage real-time recognition on computers having 4Gb of RAM, a dual core processor and a SSD of 64 GB,inferior configurations haven't been tested.

3.4. Costs

The design described here is of course flexible, meaning any online service could be used alternatively to Microsoft's. The calculation of costs for such an access control system was made considering an arbitrary a company size of a hundred workers (big company on the SME environment). Considering each of the employees comes into the company's building twice a day every working day, it is 4000 controls a month. If those faces are all stored locally, they should be recognized and therefore not generate any cost. If every day fifty unknown person comes into the building it will make about a thousand controls that are not perceived by the local recognition system. Those numbers all falls under the "free calls pool" of Microsoft Azure subscription, even considering that some queries of the analysis must be done in several steps, generating as many calls. However, this represents a laboratory reality which always differs from the "field". For the same amount of people, used in a production environment, the price of the online analysis has been calculated to be about 10 to 15 dollars a month, taking in account all the frequent errors of the software. [12]

As a onetime cost, it is necessary to get a small computer and a webcam to run the software. Multiple devices have been assessed on that purpose, all in a price range of 250 to 350 dollars for the computer and as for the camera between 50 and 80 dollars. For a total cost of 300-430 dollars a door. Counting the cost of electricity to power the system, the total cost of the installation is estimated to 1500 dollars for a period of 5 years (total cost of ownership).

4. Reliability

The test program realized to proof this design is able to distinguish similar faces such as twins easily. The confidence criterion has been configured severely, to make sure the local recognition system wouldn't give any false positive. This confidence has been set according to previous researches. [13] Tests have been repeated several times on thousands of frames without any mistake from the software.

To assess the efficiency of the software, some further tests and comparison have been conducted. The computer has been presented pictures from five pairs of twins identified in the database and two pairs of pictures of the same person on different pictures and has to differentiate them. Humans, on the other hand, have been presented a similar set of pictures and were simply asked, having two seconds for each picture to tell which subjects were twins and which were not. [14]

The precision of the software couldn't be assessed with accuracy, as so far, the program hasn't been cheated on successfully. Whenever the confidence of the local face analyser is too low, faces are

sent online for analysis. Since the program is at its final stage of development, the success rate has been of a hundred percent. Therefore, the upcoming paragraph, assessing the reliability of such systems, is based on external information and other systems.

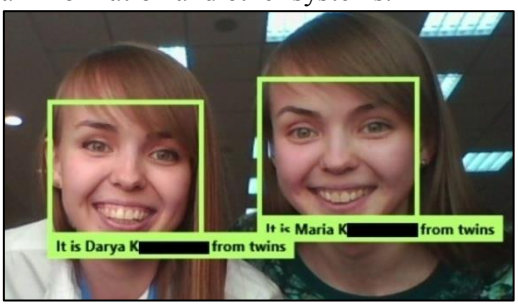

Figure 4. Twins differentiation.

Figure 5. The common test sample between the computer and humans for twins' assessment.

Table 2. Software control.

Feature	Value
Couple tested	7
Total frames	2500+
Accuracy	100%

Table 3. Human control.

Feature	Value
Human subjects	18
Couple tested	6
Total frames	72
Accuracy	61%

Unveiling its last iPhone, Apple claimed its face recognition system has a reliability of one in a million, meaning that once in a million times two faces would be confused and recognised as being the same, this is the closest comparison possible to do to the online services used. [14] To correlate this number we can take the code of a credit card in Russia, 4 digits or 10'000 possibilities, fingerprints , reputable unreliable once in 50'000 samples or an average home key (6 tumblers, 7 heights), which makes about 120'000 possibilities. Weather the reliability of the system is comparable to Apple's claim about its owns is discussable, but, nevertheless, the tests in laboratory are in favour of assessing a very high index of reliability forcomparable face recognition systems.

5. Conclusion

As underlined in this presentation, face recognition is a fast-developing market at the moment, much is already done but much is left to be built and this design has a place in the development of the market. Every major actor involved in security should now consider getting themselves an access to this kind of technology, especially now it is more accessible than ever and as the market trend makes it very profitable.

In the future, the detection will be improved by assessing the liveness of faces. Checking that we are not given a picture of a face but that it is a genuine face we have in front of the camera. This can be made by different methods, but the most adapted to a system of those dimensions is the analysis of the micro-behaviour of the eyes. [16-19] The identification system on RAM will also be compared in efficiency to a YOLO system (You Only Look Once) in order to asset their respective efficiency and choose the most appropriate technology to keep a target acquired and analyse it only once.

This kind of system could also be used on security cameras to get frames with a higher resolution and filter them threw an artificial intelligence able to understand which frames are relevant by an analysis of the pictures. Allowing selecting only relevant frames for storage, gives the possibility to significantly augment the quality of the camera's captors without being confronted to the problem of the storage space saturation. Emotion recognition and specifically this design can be adapted to the numerous of other uses such as home automations, alarms, research of wanted persons and many others that haven't been mentioned in this article. It is up for everyone, on this new market, to develop their own ideas.

Of course, this paper wasn't about a purely technical breakthrough, however I hope that the reader understands better now the face recognition market, how to use it efficiently and make it profitable, in particular with the design offered.This kind of design will make the difference between an emerging market and a fully grown and accessible one, bringing a new technology to the consumer. In other words, I want everyone to understand how face recognition systems are now in the reach of their hands.

6. References

[1] *Digital Photography review* (Access mode: https://www.dpreview.com/articles/5778663183/ten-unique-cameras-from-the-dawn-of-consumer-digital-photography) (20.8.2013)

[2] Philips P J, Flynn P J, Scruggs T, Bowyer K W, Chang J, Hoffman K, Marques J, Min J, Worek W 2005 Overview of the face recognition grand challenge, Computer Vision and Pattern Recognition *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* DOI: 10.1109/CVPR.2005.268

[3] Zhao W, Chellappa R, Philips P J, Rosenfeld A 2003 Face recognition: A literature survey *ACM Computing Surveys* **35** 399-458

[4] Gates K A 2011 *Our biometric future: facial recognition technology and the culture of surveillance* (New York University press) p 263

[5] Rybintsev A V, Konushin V S and Konushin A S 2015 Consecutive gender and age classification from facial images based on ranked local binary patterns *Computer Optics* **39(5)** 762-769 DOI: 10.18287/0134-2452-2015-39-5-762-769

[6] Nikitin M Yu, Konushin V S and Konushin A S 2017 Neural network model for video-based face recognition with frames quality assessment *Computer Optics* **41(5)** 732-742 DOI: 10.18287/2412-6179-2017-41-5-732-742

[7] Protsenko V I, Kazanskiy N L and Serafimovich P G 2015 Real-time analysis of parameters of multiple object detection systems *Computer Optics* **39(4)** 582-591 DOI: 10.18287/0134-2452-2015-39-4-582-591

[8] Jaiswal S, Bhadauria S S and Jadon R S 2011 Comparison between face recognition algorithm Eigenfaces, Fisherfaces and Elastic Bunch Graph Matching *Journal of Global Research in Computer Science* **2(7)** 187-193

[9] Yang M-H 2002 Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition* 215-220

[10] Turk M A and Pentland A O 2002 *Face recognition using Eigenface* (The Media Laboratory MIT)
[11] Kalinovskii I A and Spitsyn V G 2017 Review and testing of frontal face detectors *Computer Optics* **40(1)** 99-111 DOI: 10.18287/2412-6179-2016-40-1-99-111
[12] *Microsoft's Computer Vision API Version 2.0 documentation, Microsoft* (Access mode: https://docs.microsoft.com/en-us/azure/cognitive-services/computer-vision/home) (22.8.2018)
[13] Vizilter Yu V, Gorbatsevich V S, Vorotnikov A V and Kostromov N A 2017 Real-time face identification via CNN and boosted hashing forest *Computer Optics* **41(2)** 254-265 DOI: 10.18287/2412-6179-2017-41-2-254-265
[14] *How secure is Face ID?* (Access mode: https://www.macworld.co.uk/feature/iphone/how-secure-is-face-id-3663992/) (01.11.2018)
[15] Z Caplova, Obertov Z, Gibelli D M, Mazzarelli D, Fracasso T, Vanezis P, Sforza C and Cattaneo C 2017 The Reliability of Facial Recognition of Deceased Persons on Photographs *Journal of Forensic Sciences* **62** 1286-1291
[16] Pan G, Wu Z and Sun L 2008 Liveness detection for face recognition, recent advances in face recognition *IntechOpen* 9 DOI: 10.5772/6397
[17] Blanz V and Vetter T 2003 Face recognition based on fitting a 3D morphable model *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25(9)**
[18] Kosinski M and Wang Y 2018 Deep neural networks are more accurate than humans at detecting sexual orientation from facial images *Journal of Personality and Social Psychology* **114(2)** 246-257
[19] Pan G, Sun L and Wu Z 2017 Eyeblink-based Anti-Spoofing in Face Recognition from a Generic Webcamera *IEEE 11th International Conference on Computer Vision* DOI: 10.1109/ICCV.2007.4409068