An online real-time face recognition system for police purposes

Christos Bouras, Evangelos Michos
Computer Engineering and Informatics Department
University of Patras
Patras, Greece
bouras@upatras.gr,
emichos@ceid.upatras.gr

***Abstract*—In general, face recognition algorithms are often utilized in order to identify or authenticate a human through a camera feed. However, face recognition algorithms must adapt to real-life uncertainties and noisy environments. This work targets upon extending the existing work of facial recognition algorithms by proposing an online platform that can be used from the police forces towards effective real-time human recognition. The approach on facial recognition is to use and extend the Haar Cascade algorithm for real-time purposes, which is widely considered one for the most efficient and used algorithms for that cause. As for the criminal identification, it will made possible through image pattern recognition between the provided criminal's image and snapshots of identified faces from the livestream feed. The platform will include a live feed section with different options for video filters, enabling the user to select the best filter, depending on relevant situation of the physical surroundings for better recognition.**

Keywords—Face recognition, Haar Cascade, Python, OpenCV

I. INTRODUCTION

According to [1], Computer Vision can be defined as "an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos". Computer Vision targets at firstly understanding and then automating tasks that the human visual system can do. This field is comprised of multiple tasks, like: acquiring digital images, processing digital images, analyzing digital images, understanding analyzed images, extracting useful data. This field includes multiple well-known sub-categories, like: a) scene reconstruction, b) event detection, c) object tracking, d) learning, e) motion detection, f) visual serving, g) image restoration and many more. All these tasks effectively contribute towards creating either numerical or symbolic information. Of course, the input can be any kind, ranging from short videos, multiple camera feeds and even multi-dimensional data (like a 3-D scanner).

Today, face recognition algorithms are often utilized in order to identify or authenticate a human through a camera feed. Accuracy and efficiency of such algorithms are becoming critical prerequisites for applying face recognition within a variety of application domains like surveillance, e-banking, e-government. However, face recognition algorithms must adapt to real-life uncertainties and noisy environments. In today's world, face recognition is more and more utilized in order to search and identify shoplifters, retail criminals, or people with a history of frauds or crimes. This means that their pictures can, after being processed, be matched against a large criminal database and prevent and reduce crime rates by identifying them from installed cameras. The ability to efficiently recognize individuals through a combination of their facial characteristics is certainly an important phenomenon and the goal of facial recognition is something already apparent on our everyday lives.

For face recognition purposes, there exist multiple algorithms dedicated to efficient and accurate recognition. These approaches rely on feature extraction and fall under a) Geometric-based feature extraction and b) Appearance-based feature extraction. Geometric-based feature extraction are mainly based on Active Shape Model (ASM) and Active Appearance Model (AAM), which were deeply studied, tested and evaluated by the authors of [2]. Regarding ASM, many studies have been conducted towards improving the face models and edge enhancements (see [3], [4], [5]), whereas regarding enhancing the AAM approach, the improvements involve better landmark classifications in the face regions and automatically adjustable 2-D faces (see [6], [7], [8]). Moving over to the Appearance-based feature extraction approaches, these techniques are usually performed using color-based imaged and appearance-based classification and the algorithms involved are the Haar Cascade and the Local Binary Pattern (LBP). The Haar Cascade algorithm is one of the most used approaches towards face recognition and there exists a large research interest over extending its classifiers' efficiency for facial features recognition (see [10], [11], [12]). Last but not least, The LBP approach was introduced to the research community from the authors of [13] and the most important enhancements proposed so far are focused on increasing the pre-processing speeds and feature matching (see [14], [15]).

To this end, we propose an online platform that can be used from the police forces towards effective real-time human recognition. The main features of the platform will include a) inserting, editing and deleting user and criminal information and b) searching for criminals based on their picture through a livestream camera feed and identify them. The approach on facial recognition is to use and extend the Haar Cascade algorithm for real-time purposes, which is widely considered one for the most efficient and used algorithms for that cause. As for the criminal identification, it will made possible through image pattern recognition between the provided criminal's image and snapshots of identified faces from the livestream feed. The platform will include a live feed section with different options for video filters, enabling the user to select the best filter, depending on relevant situation of the physical surroundings for better recognition.

The rest of this work is organized as follows: In the following section, we demonstrate our approach during the design phase of the platform and Section III reveals the underlying architecture of the platform. Section IV relates to the explanation of the platform's features through screenshots and Section V includes the platform's evaluation and our main findings. Section VI reveals the limitations and concerns

regarding this platform and finally, Section VIII discusses the conclusion of our contribution.

II. Designing the Platform

Our goal is that our platform is at its core a human-centered approach. This means that we must: a) think about what people want to do rather than what the technology can do, b) design new ways to connect people with people, c) involve people in the design process, and d) design for diversity and be inclusive. Our system requirements can be categorized them into two different categories, functional and non-functional requirements, as depicted in Table. I. The functional requirements refer to want our platform must have, whereas the non-functional requirements refer to a quality that our platform should have and are both crucial factors regarding the acceptability and usability of our final product.

TABLE I. System Requirements

Functional	Non-Functional
• MySQL Database • Online Server • IP Cameras	• Responsiveness in other media • Change website and profile settings • Credentials authorization • Livestream support • Edit user/criminal records

For us to fully understand what future users might need from the platform and what challenges they might face, we decided to conduct a semi-structured interview, because we strongly believed that even though pre-prepared questions could easily apply, new topics could arise during the conversation with the stakeholders and we should be in a position to build upon such new topics without being too strongly attached to the questions drafted beforehand. During the interview, we made the questions to the participant and asked for permission to record the sessions. The discussion started with a set of pre-defined questions and then we were free to let the conversation go. For example, if we considered that the interviewee mentioned something important, we asked for them to tell more about that (of course no interviews were held with any kind of criminals).

The conversation with each participant helped us understand the needs, practices, concerns, preferences and attitudes of the people who might interact with our system. The most important finding from the interview was the reasons why the interviewed stakeholders would not trust state-of-the-art technologies for their field. The majority of the stakeholders stated that they do not believe that the platform would be as effective as presented to them and that such approaches can only work in highly technologically developed countries. Some other concerns raised were that people themselves are not good with computers (so any kind of new technology would greatly make their job even harder), they do not like new technologies and that it would be very difficult to get all local people to agree in this interconnected approach with the envisioned camera network.

Using all the information gathered so far, we created a draft version of the platform developed at this phase of the iterative design process. A Heuristic Evaluation (HE) was then conducted in order to evaluate the draft version of the platform. HE is a usability engineering method that requires a small set of evaluators to examine an interface according to its compliance with Jakob Nielsen's ten usability principles, or heuristics (see [16]). As the designers of the platform, the merged information received from the evaluators provided us with a weighted list of findings (according to their severity), alongside with propositions on how to resolve the issues mentioned. Thus, we were able to analyze the draft platform's usability, functionality and design in a standardized structure. We enumerated the findings from the study of the HCI evaluators gathered, which offered their insights and recommendations to enhance the user's experience (UX).

III. Platform Architecture

We suggest a complete systemic approach through a crime detection platform to be used in police headquarter/precincts. The platform considers the Haar Cascade algorithm for the facial recognition detection and extends it by performing real-time recognition from the connected cameras in the system. The platform supports two different types of users, namely a) police employees in the headquarters/precincts and b) police administrators, with a higher level of access and also responsible for database maintenance. The main features of the platform include:

- managing (create/update/delete) employees' information
- managing (create/update/delete) criminals' information
- applying six (6) different filters in the live feed cameras
- searching for criminals based on their database picture through the live feed cameras and identifying them

As for the criminal identification, it is made possible through image pattern recognition between the provided criminal's image and snapshots of identified faces from the livestream feed. The live feed facial recognition approach is explained in Fig. 1 below. The technical specifications for the livestream platform are listed in Table. II, offering adequate information on the tools that made this platform come to life.

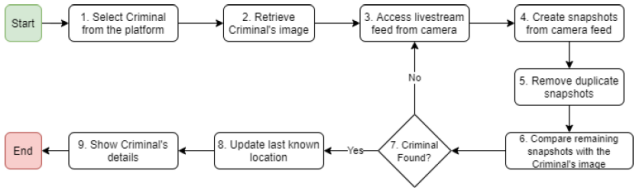

Fig. 1. The live feed facial recognition approach.

TABLE II. Technical Specifications

Integrated Development Environment (IDE)	InteliJ PyCharm 2020.3.5 Community Version
Core programming language	Python 3.6.0
Secondary programming languages	HyperText Markup Language (HTML), Cascading Style Sheets (CSS), Boostrap 4.0.0
WEB framework	Flask 1.1.2
Computer Vision Library	OpenCV
Relational Database Management System (RDBMS)	MySQL (Xampp)
Server Type	Apache HTTP
IP Camera Models	(x2) Xiaomi IP Camera Mi Home Security Camera 1080p

The database has been constructed by using three different tables, the ‘users’ table, the ‘criminals’ table and the ‘contact’ table (see Fig. 2).

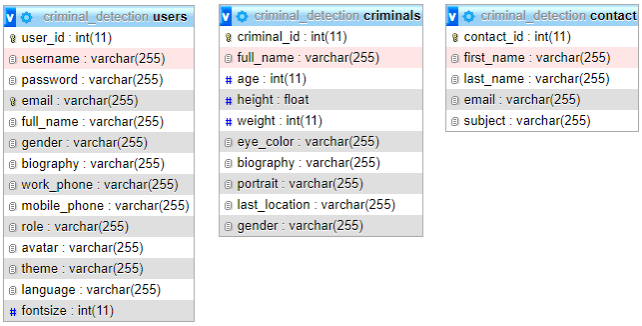

Fig. 2. The SQL database structure

Due to re-designing the website, for the different languages in the website and in order for the website internationalization to be efficient in terms of coding, we created two different Python files, called EnglishLanguage.py and GreekLanguage.py. These files serve as ‘translation libraries’, storing the translated variables under the respective Python classes. These greatly helps designing the website, as now there is now need to create each webpage two times and hardcode all information in the HTML file, but we can simply include the equivalent Python file (English or Greek) where needed and access the variables after created the respective class initializations through Python objects. This also helps our website in the distant future, as in order to insert a new language in the webpage, it is not necessary to re-design anything and the only thing needed is a translation on the respective Python file!

In order to implement streaming in Python using the Flask framework, we must use Flask’s native support for streaming responses through the generator functions, which are basically special functions that can be interrupted and resumed at any time. A generator function is able to return multiple results in a sequence, thus Flask takes this into advantage through generator functions in order to implement the streaming features. An extension of this is the use of multipart Responses. Streaming can be uses in order to have each chunk replace the previous one in the page, enabling streams to play or animate in a browser. This technique enables each chunk in the stream to be simply an image, and then we get the whole video live-stream in the browser through repetitive chunks of images and in order to do that, Flask uses the multipart responses. Multipart responses consist of headers that include one of the multipart content types and are followed by the parts, separated by a boundary marker and each having its own part specific content type. For the purposes of our platform, we need each new chunk to replace the previous one as single-instance images, thus we will use the ‘multipart/x-mixed-replace’ content.

Upon concluding on how the streaming services will operate, the next issue is deciding on the face recognition library. We used the Open Source Computer Vision (OpenCV) library, which is a real-time cross-platform library and free for use under the open-source Apache 2 License [17]. OpenCV is written is C++, its primary interface is in C++ and All of the new developments and algorithms appear in the C++ interface, yet there exist bindings in different programming languages and the Python language (which we selected for our platform) is one of them. OpenCV was built to provide a common infrastructure for Computer Vision applications and to accelerate the use of machine perception in the commercial products. As of today, the library has more than 2500 optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc.

IV. The Platform in Action

When launching the website for the first time, we are greeted with the welcoming page of the platform. As stated in previous sections, login credentials are required to enter the platform. Upon correctly entering our user credentials, we can now view the main page (Home page) of the website. As we can see below, we are introduced with relevant information on what we can do in the platform and we can access the platform’s features either from the hyperlinks in the ‘Features’ section or from the website header (see Fig. 3). Now, let’s say that we want to access the management panel for the criminals. By keeping conformity with these colors, we will help users understand the type of action that takes place, as they can mnemonically link colors with actions. As one may observe in the figures below, the header color of the pop-up and the relevant submit action button are colored the same, exactly following the aforementioned approach (see Fig. 4 and Fig. 5).

The Live Feed section offers six different options for video filters, enabling the user to select the best filter, depending on the relevant situation of the physical surroundings. Again, each time a criminal is recognized, a snapshot is taken directly from that camera feed. The six (6) different filters that can be used for face recognition and all snapshot are categorized in the specific folder created for that criminal, where snapshots are saved inside the folder (the filename specifies the datetime and the filter used by the camera) and re-categorized depending on the camera that took the snapshot. This helps us, as we can retrieve his last known location by knowing the location of the camera that took the screenshot (see Fig. 6 – Fig. 8).

The Settings page can be found on the top-right corner of the website, under the welcoming message to the user. In the final version of the platform, we added a ‘My Settings’ section on the top-right corner of the website. The header of the website now includes a welcoming message to the user and by clicking on it, users can select the ‘My Settings’ section or they can simply log out of the platform. In order to help user reduce the amount of time needed for settings tasks, the ‘My Settings’ and ‘My Profile’ sections appear inside the same webpage. The ‘My Settings’ tab allows users to change the website’s appearance (e.g. light/dark theme, different font sizes and English/Greek language), whereas the ‘My Profile’ tab enables users to change their account settings.

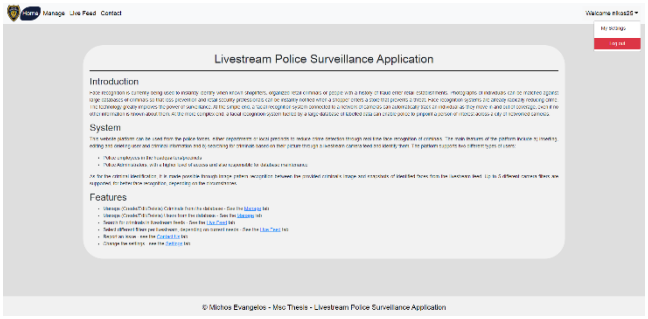

Fig. 3. The Home page

Fig. 4. The Manage Criminals page

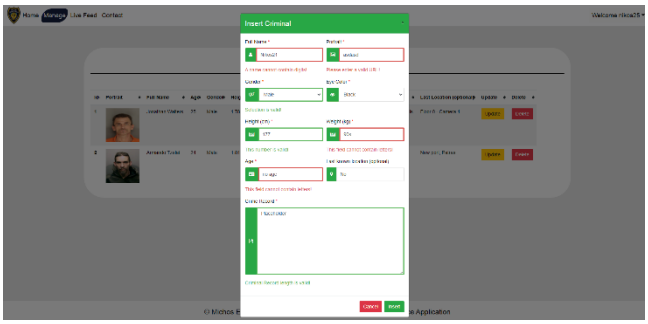

Fig. 5. Keeping the color conformity based on the button colors.

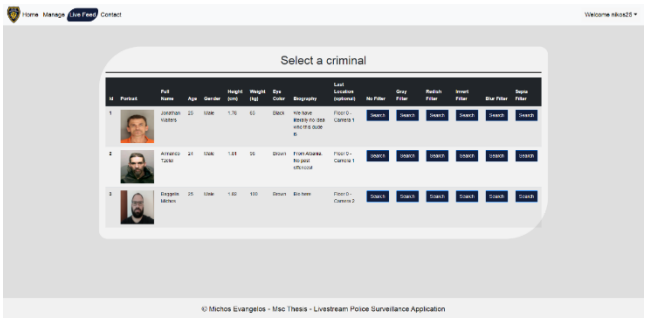

Fig. 6. Selecting a criminal for real-time recognition

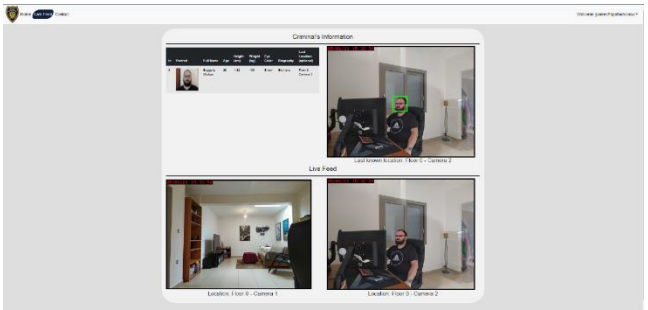

Fig. 7. The Live Feed page

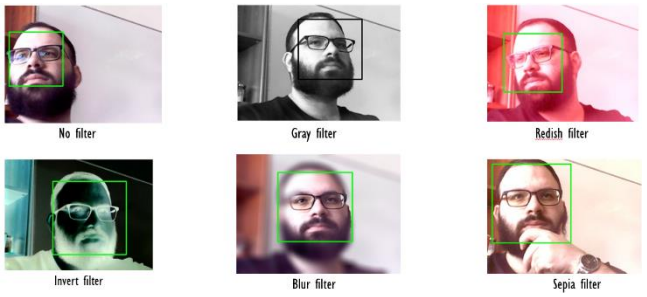

Fig. 8. The different applicable filters in the Live Feed page (author's images).

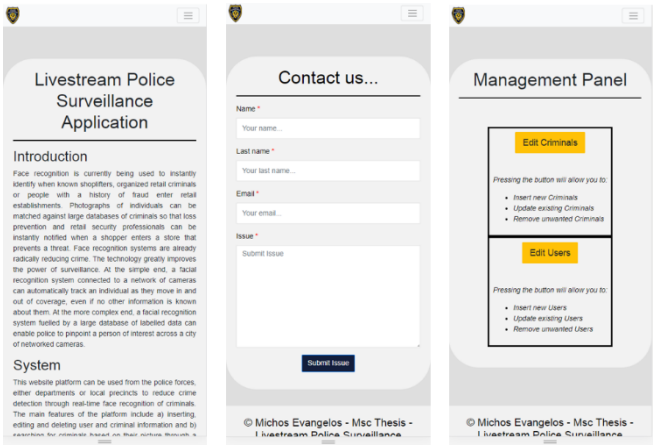

Fig. 9. Mobile Responsiveness

V. Platform Evaluation

In order for the evaluation procedure to be valid and reliable, the platform had to be tested from police officers from a local police department. To extract data from the evaluation process and mostly specific information for the characteristics of the system, three different types of questions were used. The first one is based on close type questions in which the respondent has to choose between "YES" or "NO" as potential answers regarding general questions about the system and the user's interaction with the platform. The next questions involve answers based on Likert scale that starts from 1 (Strongly Disagree) and ends with 5 (Strongly Agree) regarding how user experience (UX) on the website. The last section ends with some open-ended questions in which respondents can freely describe their overall impressions of the platform. The purpose of the work was not to make a detailed statistical analysis, but to use the questionnaire as an additional mean of evaluation. After all, the number of participants was relatively small and equal to 30, by the time the submission deadline was closed. As a result, the following will tackle the most important and interesting questions of the questionnaire, but the sample is not adequate for further statistical analysis.

The reason why we used Likert scale (which is one of the most widely used question types in a survey) in the questionnaire for the specific questions regarding the platform design is as follows. A Likert scale question is a psychometric scale where questions based on this scale are used in a survey. In a Likert scale survey, respondents don't choose just between 'YES' or 'NO', but there are specific choices based on 'agreeing' or 'disagreeing' on a particular survey question. Likert scale survey questions are essential in measuring a respondent's opinion or attitude towards a given subject and is an integral part of market research. Likert scale is typically a

five, seven, or nine-point agreement scale used to measure respondents' agreement with various statements. In general, a series of statements each designed to view a construct from a slightly different perspective is leveraged. This technique's power is that it works across disciplines-it is just as applicable to a social science construct as it is a marketing one. Likert scale questionnaires offer the following advantages:

- Easy to design and apply on the survey
- Can include items not related to expression
- Provide a range of opinions to the respondents
- Produces accurate and quality information, reducing measurement error
- Allows to perform the necessary analysis to achieve the research objective
- Can be compared with previous service evaluations

To begin with, regarding the introductory demographics (see an indicative example in Fig. 10), the different ages of the stakeholders involve a 30% responses from stakeholders aged between 40-49, 23.3% aged between 20-29 and 23.3% aged between 50-59 years old. Regarding the years employed in the precinct, 23,3% of the respondents worked 10-19 years, 20% worked 20-29 years and 16.7% worked 7-9 years.

Regarding the most important usability evaluation questions (see an indicative example in Fig. 11), results revealed the vast majority of the employees stated that the purpose and scope of the website was clear (86.7%), whereas results revealed that 90% of them stated that they felt that the website endorsed interactivity. Regarding the website's responsiveness in different platforms, 56.7% of them stated that they tested the website with such media (mostly mobile phones), whereas 36.7% of them did not (even) check. Regarding the language support, the hopeful amount of 76.7% stated that the platform could be understood from people of different nationality, meaning that they found the setting to change the language in the settings. On the other hand, the 23.3% of them either wanted to be in a neutral position or did not understand that they could change the language.

Next, we will look at the specific usability evaluation questions on the website's design and interactivity. In order to save up space, since the usability questions could be answered from Strongly Disagree -> Strongly Agree, the responses can be merged, so that a 'General Disagreement' refers to merging both the 'Strongly Disagree' and 'Disagree' responses, whereas the 'General Agreement' refers to merging the percentages of 'Agree' and 'Strongly Agree' responses. To this end, the responses from the website's evaluation questions are presented below in Table III.

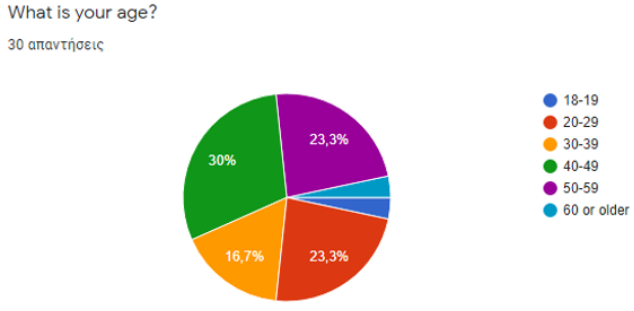

Fig. 10. An indicative demographics question

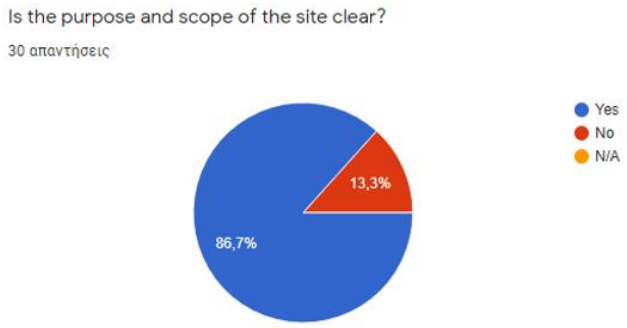

Fig. 11. An indicative evaluation question

TABLE III. WEBSITE'S DESIGN EVALUATION

Usability Question	Response
Is the management panel appealing?	General Agreement 66.7 %
Are there any choices for recovering from errors?	General Agreement 66.7 %
In overall, are you satisfied with how easy it is to use this website?	General Agreement 80 %
Does this site help you to complete your work efficiently?	General Agreement 80 %
Do you believe you will eventually become productive in your work by using this site?	Neutral 50%
Does the site offer innovative features for its field?	General Agreement 53.3 %

The last questions in the questionnaire were open type questions, where employees could openly express their honest opinions on the platform without restrictions. On whether they would continue to use the website on a regular basis, 14 of the 22 submissions were positive, while the remaining 8 of them where still not convinced about the platform due to the country it is being implemented mostly (Greece). Regarding the question on what site feature they would like to see most improved, the responses (12 in total) vary between users that were happy with the design and could not think of something to add and those that very vaguely suggested changes like in the home page, or another way of searching for criminal in the livestream page. Regarding the question of what changes they would generally suggest for the website, we received 13 responses and the requests varied between larger databases for criminals and more recognition features (license plate, human pose and stolen clothes), while once again, others had nothing to note. On the question of whether this platform could effectively help the fight towards crime, almost 9 out of the 13 responses (69%) responded positively where the rest 4 responses (31%) were still not convinced that such a platform could work (the country plays a very important factor on this). Lastly, we asked employees to use only one word to rate the whole platform and to this end, 17 out of the 18 response (94%) were positive, using words like insightful, appealing, promising, innovative, useful and satisfying, whereas only one person (6%) judged the website as complicated.

VI. LIMITATIONS AND CONCERNS

So far, our approach could seem ideal for the interested parties. Yet, significant limitations and concerns exist when facial recognition applies for criminal identification. The very first fear that should be generated in everyone's head is the fact that it may not work well enough. To be fair, this is indeed an issue, since technology is always prone to false positives and false negatives and particularly when utilized with noise

imagery from the Close-circuit television (CCTV) cameras that have been already installed years or even decades ago and the state has never considered updating the camera infrastructure in the city. And this may not have been a great issue so far, but when face recognition cameras begin to be used in order to arrest, convict or imprison people, then a possibly faulty basis can certainly cause real harm.

What is even worse is the fact that errors that occur are not evenly distributed, as facial recognition systems have been regularly found in the past to be inaccurate at identifying people with darker skin [18]. This can of course be solved with a new camera-system infrastructure, yet this demands an adequate amount of funds from the state, set aside the fact that as technology advances are facial recognition becomes more and more included in our everyday lives, people might eventually think that facial recognition inherently undermines freedom by enabling perfect surveillance of everyone, all the time.

Furthermore, it has been revealed that even though police tend to praise the different technologies used in order to improve future investigations, many agencies prefer to keep their methods as secret as possible [18]. This results in people getting arrested or convicted without them knowing they were recognized through face identification from cameras and exactly because the police does not treat facial recognition as adequate evidence in court, this approach sometimes does not turn up in public documents and has not been many times a subject of judicial rulings.

Last but not least, people know very little about facial recognition technology, set aside how efficient it will be in reducing crime rates. When systemic federal regulations or permitting processes are not apparent, police can only gather information from stories, interviews, public reports and investigative reports. This could mean that even though a police department or a police precinct may publicly acknowledge the fact that facial recognition is utilized from their side for criminal apprehension, they might not collect or share any king of tangible metrics or analytics on how effective is their approach, leaving everyone on the outside.

VII. Conclusion

In today's world, face recognition is more and more utilized in order to search and identify shoplifters, retail criminals, or people with a history of frauds or crimes. This means that their pictures can, after being processed, be matched against a large criminal database and prevent and reduce crime rates by identifying them from installed cameras. To this end, this contribution targets upon extending the existing work of facial recognition algorithms by proposing an online platform that can be used from the police forces towards effective real-time human recognition. The main features of the platform will include a) inserting, editing and deleting user and criminal information and b) searching for criminals based on their picture through a livestream camera feed and identify them.

References

[1] https://en.wikipedia.org/wiki/Computer_vision

[2] M. Iqtait, and F. Mohamad, "Feature extraction for face recognition via Active Shape Model (ASM) and Active Appearance Model (AAM) ", IOP Conference Series Materials Science and Engineering. 2018.

[3] T. Le, and V. Truong, "Face Alignment Using Active Shape Model And Support Vector Machine". International Journal of Biometric and Bioinformatics, 2011.

[4] B. van Ginneken, A. F. Frangi, J. J. Staal, B. M. ter Haar Romeny and M. A. Viergever, "Active shape model segmentation with optimal features", in IEEE Transactions on Medical Imaging, 21(8), pp. 924-933, 2002.

[5] H. Lu, and F. Yang, "Active Shape Model and its Application to Face Alignment", in Subspace Methods for Pattern Recognition in Intelligent Environment. Studies in Computational Intelligence, vol 552. Springer, Berlin, Heidelberg, 2014.

[6] T.F. Cootes, and C.J. Taylor, "Statistical models of appearance for computer vision". 2004.

[7] T.F. Cootes, G.J. Edwards, and C. Taylor, "Active Appearance Models. Pattern Analysis and Machine Intelligence", IEEE Transactions, pp. 681 – 685, 2001.

[8] G. Tzimiropoulos, and M. Pantic, "Fast Algorithms for Fitting Active Appearance Models to Unconstrained Images", International Journal of Computer Vision, 122, pp. 17–33, 2017.

[9] X. Zhao, S. Shan, X. Chain, and X. Chen, "Locality-Constrained Active Appearance Model", in Computer Vision – ACCV 2012. Lecture Notes in Computer Science, vol 7724. Springer, Berlin, Heidelberg, 2013.

[10] L. Cuimei, Q. Zhiliang, J. Nan and W. Jianhua, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers", 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), pp. 483-487, 2017.

[11] R. Padilla, C. Filho, and M. Costa, "Evaluation of Haar Cascade Classifiers for Face Detection", 2012.

[12] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features", IEEE Conference Computer Vision Pattern Recognition. 1. I-511, 2001.

[13] T. Ojala, M. Pietikainen and D. Harwood, "A comparative study of texture measures with classification based on feature distributions", Pattern Recognition vol. 29, 1996.

[14] A. Gaikwad, "LBP and PCA based on Face Recognition System", 5, pp. 368-373, 2018.

[15] D. Huang, C. Shan, M. Ardabilian, Y. Wang and L. Chen, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey", in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 41(6), pp. 765-781, 2011.

[16] https://www.nngroup.com/articles/ten-usability-heuristics/

[17] https://opencv.org/about/

[18] J. Schuppe, "How facial recognition became a routine policing tool in America", NBC News. Retrieved May 2021.