Low-Resource Face Spoofing Detection in a User-friendly Manner.

Rien Heuver
University of Twente
P.O. Box 217, 7500AE Enschede
The Netherlands
rienheuver@gmail.com

ABSTRACT

This research aims to distinguish real faces from spoofed faces when presented to a mobile phone front-camera. Since face recognition is potentially an incredibly user-friendly method to authenticate a user, it is important that the false acceptance rate (FAR) and false rejection rate (FRR) are as low as possible. However, the detection method should not require a large amount of time to classify nor should it depend on resources not present on regular mobile phones. Current research proves that texture analysis through Local Binary Patterns (LBP) is a promising method of distinguishing real from fake faces. This research creates a new small dataset consisting of images taken with a mobile phone front-camera. The calculated LBP's are used to train a machine learning binary linear classifier. A FAR of 2.5% and a FRR of 7.1% were found with an average computation time of 67 milliseconds. These results indicate that an effective low-resource anti-spoofing method is possible.

1. INTRODUCTION

In modern day technology, speed and user-friendliness are two key factors for successful software. The popularity of mobile phones is ever increasing, but they are still limited in their capacity. To identify people automatically, we can use biometric techniques, such as face recognition [7], effectively. However, most of these identification processes can be spoofed easily [1]. Many solutions to detect the liveliness of the user, in order to detect such spoofing attacks, otherwise known as presentation attacks, are available [3], though their practicality can be questioned. An example is to ask the user to perform a certain task, which can then be checked by the system. If the task is performed correctly the user is accepted to not be fake. This is called a challenge-response technique [4]. Such a technique, however, is not user-friendly at all. Another approach to distinguishing a fake user from a real one is by measuring the temperature of the user, thus clearly proving its liveliness. However, this requires technology that is not found on current mobile devices, thus it is not an acceptable solution either. The research described in this paper looks for a low-resource and user-friendly approach to prevent such spoofing attacks.

A user trying to spoof a face recognition system has a variety of choices to do so. He can present a still image, he can present a video or he can present himself with a 3D-mask of the user he wants to imitate. Since the latter is a complicated technique to perform for an attacker, the focus of the method in this research is on detecting still images and video attacks. When an attacker presents a picture of the user or a video of the user, we call this a 'recaptured image'. A promising method is micro-texture analysis, which focuses on the minor difference between an actual face and a recaptured image. Recaptured images of faces, such as regular prints, usually contain small defects that can be detected through texture analysis. This method uses only one frame to detect whether or not the presented user is real, thus it automatically also works for video spoofing attacks. The analysis method implemented is called Local Binary patterns (LBP); a technique that examines neighbouring pixels to determine patterns in images. A Support Vector Machine (SVM) is then trained with the LBP's of a set of images. When a new image is then analysed and run through the SVM, it should be correctly classified as a real presentation or a presentation attack.

In the following parts, this research will be elaborated. In section 2 the addressed problem will be further explained. Section 3 will then further explain the research question. In section 4 related work will be discussed and possible improvements will be proposed. Further explanation of the methods and approaches of this research used will be discussed in section 5. Results will then be shown in section 6. Discussion on the research will be found in section 7.

2. PROBLEM STATEMENT

A high variety of spoofing attack detections exists. However, in this research we aim to provide a method that users would want, or at least would not mind, to use to unlock their mobile phones. Therefore it is assumed that the unlocking mechanism is used frequently and thus plenty of successful spoofing detection methods are no longer wanted. A challenge-response mechanism for example would require too much effort and time from the user if it would be used frequently throughout the day. In order to find a method that is less intrusive and fast performing, it is necessary that the found method relies only on available hard- and software in current day mobile phones. Therefore, using face recognition was chosen as an assumed working identification method.

In the current state of the art, various techniques have been researched and found to prevent spoofing attacks on face recognition. However, they vary highly in reliability and computational time. Many practical applications of face recognition are likely to be found on mobile phones, where authorization is required or desired frequently. The

current state of the art is not conclusive about how reliable and fast spoofing prevention methods will perform on mobile devices when served with simple mobile phone captured images. The proposed research will try to find a method that fills this gap to make face recognition a more practical and reliable tool for mobile devices.

3. RESEARCH QUESTION

A mobile phone has limited camera quality, especially with most front-facing cameras. The goal is to detect a real face, one that is not recaptured, in as short a time as possible, thus to be user-friendly and truly mobile. Mobile phones are, however, limited in processing power, thus complicated techniques are not likely to be fast.

- How fast and reliably can a real face be detected with a mobile phone camera?
    - To what extent can a low false acceptance rate (FAR) and a low false rejection rate (FRR) be achieved?
    - To what extent can this be performed in a non-intrusive way and without high waiting times?

In section 4 it is explained that a FAR of 2.2% and a FRR of 13% are the current state of art. Therefore, this research aimed to achieve comparable results, where comparable is defined as a deviation of at most 10%. Thus a FAR of at most 2.4% and a FRR of at most 14.3%.

4. RELATED WORK

Määttä et al. [6] investigated face spoofing detection of single images using micro-texture analysis. Their experiments were run on photo-quality and laser-quality printed faces. However, the idea of this research was to perform experiments with images captured by the mobile phone. The experiment was also run on attacks by recaptured images of digital displays as to compare the results of those to paper-printed attacks.

Bai et al. [2] too investigated face spoofing detection on single images. They too used a texture-based analysis method. In their experiments, however, they used high quality cameras to capture the images. In this research the images were captured with a low-quality mobile phone camera. Besides that, the method by Bai et al. took almost five seconds which this research aimed to improve.

Other methods for detecting face spoofing exist too, but they either use too complex techniques (such as thermal cameras) which are not available on regular mobile phones, or techniques that require more time, for example to capture multiple frames or even request a certain response from the user (challenge-response techniques).

Regarding the related worked mentioned, this research aimed to improve testing on a more varied dataset; namely both paper recaptured faces and digital display (screen) recaptured faces. Besides that, it aimed to improve the time required to classify an image as either a real face or a spoofing attack.

5. METHODS AND APPROACH

5.1 Dataset

As explained in 4, this research aims to test on both paper recaptured spoofing attacks as well as digital screen recaptured spoofing attacks. Since this research is aimed on low-resources and mobile phones, a new dataset had to be build. Twenty people were asked to capture their face. These captures were made with the front-facing camera of the same mobile phone, thus testing on low-quality images. The phone used for this research is a Oneplus 2 from the Chinese manufacturer Oneplus. A variety of lightings and some small variety in poses were made by the twenty subjects. This resembles real-life scenarios in which the subject will want to use face-identification on his/her mobile phone. A few examples of captured images are shown in figure 1.
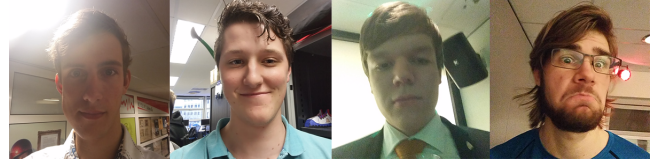

Figure 1. Example pictures taken by subjects

Since only the actual face, and not the background of the picture, is of interest to this research, the pictures were normalized. Facial recognition/identification systems will also normalize pictures; therefore this is not regarded as part of the process nor as a contribution to the computation time the provided method in this research will require. Because of the limited scope of this research, thus the pictures where normalized manually using Adobe Photoshop CC 2015. Examples of the normalized pictures in figure 1 are shown in figure 2. The normalization was done by boxing the face of one subject and taking that as the first normalized image. The distance between the eyes of that subject was then measured and used to normalize the remaining pictures, keeping the size of the box equal to the one of the first subject and the distance between the eyes equal to the distance between the eyes of the first subject.
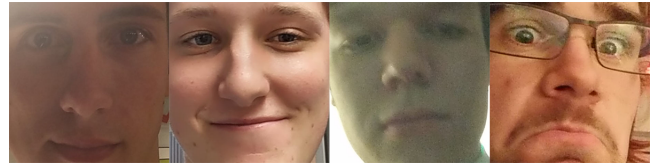

**Figure 2. Normalized versions of figure 1**

5.2 Image analysis algorithm

The method in this research implements a micro-texture analysis based method called Local Binary Patterns (LBP). The LBP-algorithm consists of the following steps:

1. Split the image up into cells

2. For each pixel in the cell, compare its value to the value of its surrounding neighbours.

3. If the neighbours value is higher, register a zero, otherwise register a one. This results in an 8-digit binary number.

4. Compute the histogram over the cell of the frequency of occurrence of each binary number.

5. Concatenate the histograms of each of the cells of the image.

The LBP-algorithm in this research splits the images up into four cells over which it calculates the histograms. An example for LBP-algorithm over a single pixel is given in
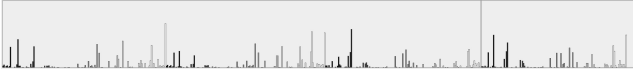
figure 3. An example histogram of one cell is given in figure 4 and an example of concatenated histograms is shown in figure 5.

58	54	55
56	55	56
58	58	59

0	1	1
0		0
0	0	0

01100000
(96)
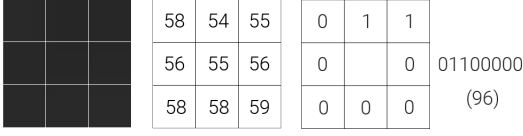
Figure 3. Steps 2 and 3 of the LBP-algorithm
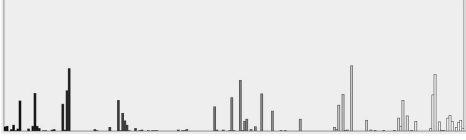
Figure 4. Step 4 of the LBP-algorithm

Figure 5. Step 5 of the LBP-algorithm

5.3 Classification of images

The histograms calculated per image (see section 5.2) can be used to classify the images, where we consider two classes: real faces and spoofed faces. This classification is done through machine learning. A certain amount of images (histograms) is labelled with the class they belong to. They are then used to train a support vector machine [5], a binary linear classifier. This classifier works as follows:

The concatenated histograms of the images are mapped to points in a space. This mapping is done in such a way that the points belong to a certain class are mapped close to another while keeping a maximum distance to the other class. Thus there will be two groups of dots in the space, one group to which the real faces belong and one to which the spoofed faces belong.

Once the classifier has been trained with some real and spoofed faces, it can be used to classify new faces. For such a face (image) we once again compute the concatenated histogram with LBP which we then feed to the classifier. The classifier will map it into the space it trained before and will identify to which group it is situated closest. The classifier can then use the label of that group as the class the analysed face belongs to.

6. RESULTS

6.1 Trained model

The dataset consists of twenty faces, twenty paper-recaptures and twenty screen-recaptures resulting in a total of sixty images. To train the classifier, six subjects, a total of eighteen images, were separated from the dataset. The remaining fourteen subjects, forty-two images, were used to generate tests. The model derived from training with six subjects is shown in figure 6.

It seems to sufficiently separate the real faces (coloured blue) from the paper and screen faces (coloured yellow).
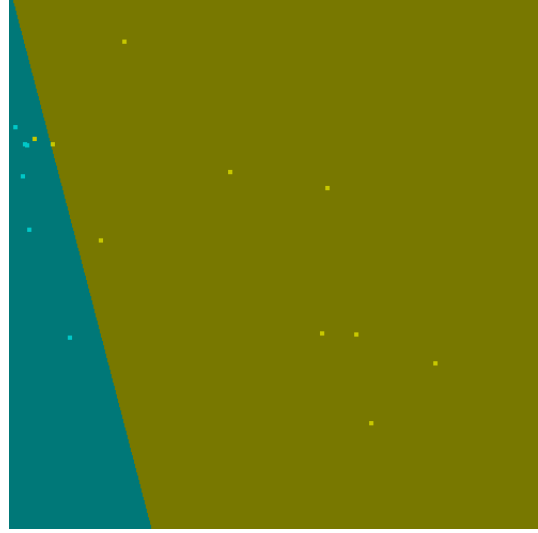
Figure 6. The linear binary classifier trained with six subjects

6.2 Testing subjects

However, when the other subjects are plotted in the same way, the results as shown in figure 7 are found. In figure 7 the screen faces have been coloured pink to show the clear difference between the two spoofing methods.
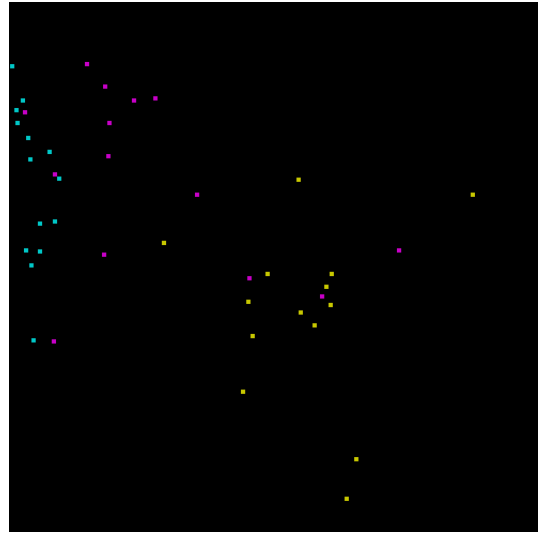
Figure 7. The plotted remaining fourteen test-subjects

As further ellaborated on in section 6.3, there was only one image that was able to generate a false accept. This image was indeed a screen recapture. Nevertheless, the image is not distinctly different from the other screen recaptures to the human eye [1].

6.3 Answering research questions

Fourteen subjects were tested on the classifier that was trained as explained in 6.1. Only one image was able to generate a false accept, and three images were able to generate a false reject. Thus the achieved FAR is 2.4% and the achieved FRR is 7.1%.

---

[1]Not all subjects allowed for their pictures to be published, which is why the false accept image is not shown

Besides achieving certain percentages regarding accepts and rejects of images, the processing time of the algorithm is also of importance to this research. The average time of the classification of all forty-two images was measured, where the entire time taken to classify an image was measured. This means both the LBP-algorithm as classification through the binary linear classifier. The average time taken was 67 milliseconds.

7. DISCUSSION

The results of this research are unfortunately limited due to limited time. It requires a rather large amount of manual work and time to construct an elaborate dataset on this matter. Because of the absence of time the used dataset was quite limited. The results found are, however, nonetheless meaningful in this research.

The first sub-question, whether or not a low FAR and low FRR could be achieved, is not entirely answered. The FAR was aimed to be below 2.2% which unfortunately did not succeed, as the research achieved a FAR of 2.4% as explained in section 6.3. However, as explained in section 6.1, this could well be because of a limited dataset and limited training of the model. The results as shown in section 6 are not conclusive because of the limited dataset. However, it does show that with a small amount of training data, most classification (93.5%) can be successful. When a larger and more varied dataset is constructed, more accurate results could be determined. In a general sense, a machine learning algorithm, such as the binary linear classifier, will perform better when trained with more data. Thus the results found in this research proof to be at least promising.

The second research question was whether or not the classification of faces could be done in a non-intrusive and relatively fast way. The method discussed in this research is clearly non-intrusive. The user only has to show his/her face to the front-camera of his/her mobile phone.

About the waiting times, the following can be said. The results found in this research were generated on a desktop machine with rather sufficient (graphics accelerated) processing power. The computer took only 67 milliseconds per image to classify it. Mobile phones that are equipped with a front-camera should have sufficient processing power to still have the classification done within reasonable times.

Therefore, both research questions are not fully answered, but the found results are promising.

8. REFERENCES

[1] A. Anjos, J. Komulainen, S. Marcel, A. Hadid, and M. Pietikäinen. Face anti-spoofing: Visual approach. In *Handbook of Biometric Anti-Spoofing*, pages 65–82. Springer, 2014.

[2] J. Bai, T.-T. Ng, X. Gao, and Y.-Q. Shi. Is physics-based liveness detection truly possible with a single image? In *Circuits and systems (ISCAS), Proceedings of 2010 IEEE international symposium on*, pages 3425–3428. IEEE, 2010.

[3] W. Bao, H. Li, N. Li, and W. Jiang. A liveness detection method for face recognition based on optical flow field. In *Image Analysis and Signal Processing, 2009. IASP 2009. International Conference on*, pages 233–236. IEEE, 2009.

[4] R. M. Bolle, J. H. Connell, and N. K. Ratha. System and method for liveness authentication using an augmented challenge/response scheme, Feb. 1 2005. US Patent 6,851,051.

[5] C.-W. Hsu, C.-C. Chang, C.-J. Lin, et al. A practical guide to support vector classification, 2003.

[6] J. Määttä, A. Hadid, and M. Pietikainen. Face spoofing detection from single images using micro-texture analysis. In *Biometrics (IJCB), 2011 international joint conference on*, pages 1–7. IEEE, 2011.

[7] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM computing surveys (CSUR)*, 35(4):399–458, 2003.