

Article

In the Eye of the Deceiver: Analyzing Eye Movements as a Cue to Deception

Diana Borza, Razvan Itu and Radu Danescu *

Department of Computer Science, Technical University of Cluj-Napoca, Memorandumului Street 28, 400114 Cluj-Napoca, Romania; diana.borza@cs.utcluj.ro (D.B.); razvan.itu@cs.utcluj.ro (R.I.)

* Correspondence: radu.danescu@cs.utcluj.ro; Tel.: +4-026-440-1456

Received: 18 September 2018; Accepted: 12 October 2018; Published: 16 October 2018



Abstract: Deceit occurs in daily life and, even from an early age, children can successfully deceive their parents. Therefore, numerous book and psychological studies have been published to help people decipher the facial cues to deceit. In this study, we tackle the problem of deceit detection by analyzing eye movements: blinks, saccades and gaze direction. Recent psychological studies have shown that the non-visual saccadic eye movement rate is higher when people lie. We propose a fast and accurate framework for eye tracking and eye movement recognition and analysis. The proposed system tracks the position of the iris, as well as the eye corners (the outer shape of the eye). Next, in an offline analysis stage, the trajectory of these eye features is analyzed in order to recognize and measure various cues which can be used as an indicator of deception: the blink rate, the gaze direction and the saccadic eye movement rate. On the task of iris center localization, the method achieves within pupil localization in 91.47% of the cases. For blink localization, we obtained an accuracy of 99.3% on the difficult EyeBlink8 dataset. In addition, we proposed a novel metric, the normalized blink rate deviation to stop deceitful behavior based on blink rate. Using this metric and a simple decision stump, the deceitful answers from the Silesian Face database were recognized with an accuracy of 96.15%.

Keywords: blink detection; deceit; eye movements; iris center

1. Introduction

Deceit, the distortion or omission of the (complete) truth, is a frequent and important aspect of human communication. However, most people, even with special training, fail to detect untruthful behavior; some studies show that their odds of succeeding are worse than chance [1]. The problem is that there is no such thing as an infallible source of deceit detection.

The traditional tools for deceit detection (i.e., the polygraph tests) are based on several responses of the autonomic nervous system (ANS)— blood pressure, breathing pattern, skin resistance, etc.—correlated with the interrogation of the suspect. No standardized interrogation protocol has been proposed, so there are main ways of administering these questions [2]: the Control Question Technique (CQT)—largely used in the United States, which aims at detecting psychological responses to the questions, and the Concealed Information Test [3] (CIT)—used in Japan, which is designed to detect concealed crime-related knowledge.

Besides these classical methods, other cues of deceit detection have been considered [4]: emblems, illustrators, micro-expressions and eye movements. The eyes are perhaps the most expressive and salient features of the human face, as they convey tremendous amounts of information: cognitive workload, (visual) attention, neurological processes, just to name a few.

Popular theories state that the position of the eyes can tell deceit; such a belief is used by one of the most popular so called personal development techniques: Neuro-Linguistic Programming (NLP) [5].

One thesis from NLP states that the gaze direction can be used as an indicator to whether a person is telling the truth or not. More specifically, this theory suggests that humans tend to move their eyes to their left when visualizing past events, and tend to look right when constructing false events. A recent study [6] tested this hypothesis and found no evidence to support the idea that the eye movement patterns can indicate lying. Despite several critics [6,7] and the little scientific evidence to support the NLP, this theory is still widely spread on the Internet and used by many NLP practitioners. However, this does not imply that other eye features cannot be used as cues to deceit detection.

In the same context, it is worth mentioning the Facial Action Coding System (FACS) [8]—an anatomical methodology, developed in the late 1970s, to describe all observable human face movements. Basically, this taxonomy breaks down facial expressions into Action Units (AUs): a contraction or relaxation of a facial movement. Since its publication, it has undergone multiple revisions and updates, and now it is often used in facial expression analysis. AUs 61 to 69 describe the eye movements: eyes (turn) left, eyes (turn) right, eyes up, eyes down.

Blinking is defined as a rapid closure followed by re-opening of the eyelids; this process is essential in spreading the tears across the eyes' surface, thus keeping them hydrated and clean. The average duration of a blink is considered to be around 100–150 ms [9], although some studies suggest longer intervals (100–400 ms) [10]. Eye closures that last longer than 1 s can be a sign of micro-sleeps: short time intervals in which the subject becomes unconscious and it is unable to respond to external stimuli. A detailed survey of the oculomotor measures which could be used to detect deceitful behavior can be found in [11].

Blinks (and micro-sleeps) have been examined in a variety of multidisciplinary tasks: deceit detection [12,13], driver drowsiness detection [14,15], human–computer interaction [16] and attention assessment [17], just to name a few.

Automatic blink detection system can be roughly classified as appearance or temporal based [18]. The first methods rely on the appearance of the eye (closed or opened) in each frame to determine the blink intervals. Temporal based methods detect blinks by analyzing the eyelid motion across video frames.

In [19], the authors propose a real-time liveness detection system based on blink analysis. The authors introduce an appearance based image feature—the *eye closity*—determined using the AdaBoost algorithm. Based on this feature, eye blinks are detected by a simple inference process in a Conditional Random field framework.

The work [15] proposes a novel driver monitoring system based on optical flow and driver's kinematics. Among other metrics, the system computes the percentage of eyelid closure over time (PERCLOSE) in order to infer the driver's state. The blinks are detected by analyzing the response of a horizontal Laplacian filter around the eyes. The authors assume that when the eyes are open, numerous vertical line segments (caused by the pupils and the eye corners) should be visible; on the other hand, when the eyes are closed only horizontal lines should be observed. The system decides on the eye state by applying a simple threshold on the value of the horizontal gradient. The value of this threshold was established heuristically, based on trial and error experiments such that the number of false positives is minimized.

In [20], the average height–width eye ratio is used to determine if the eyes are open or closed in a given frame. First, 98 facial landmarks are detected on the face using active shape models; based on the contour of the eyes, the height–width eye ratio is computed and a simple thresholding operation is used to detect eye blinks: if this measure changes from a value larger than 0.12 to a value smaller than 0.02, then a blink was detected. As a static threshold value is used, this method cannot be used for real world, “in the wild” video sequences.

In [18], blinks are detected by analyzing the vertical motions which occur around the eye region. A flock of Kanade–Lucas–Tomasi trackers are initialized into a grid around the peri-ocular region and are used to compute the motion of the cells which compose this grid. State machines analyze the variance of the detected vertical motions. The method achieves a 99% accuracy rate.

A real-time blink detector designed for very low near-infrared images is presented in [21]. Blinks are detected based on thresholded image differences inside two tracked region of interest corresponding to the right and left eye, respectively. Next, optical flow is computed in order to determine if the detected motion belongs to an eyelid closing or opening action.

Automatic detection of gaze direction implies the localization of the iris centers and determining their relative position to the eye corners or bounding rectangle.

Eye tracker have been extensively studied by the computer vision community over the last decades; in [22], the authors present an extensive eye localization and tracking survey, which reviews the existing methods the future directions that should be addressed to achieve the performance required by real world eye tracking applications. Based on the methodology used to detect and/or track the eyes, the authors identified the following eye tracking system types: shape based, appearance based and hybrid systems.

Shape based methods localize and track the iris centers based on a geometrical definition of the shape of the eye and its surrounding texture; quite often, these methods exploit the circularity of the iris and the pupil [23–26]. In [26], the iris centers are located as the locus where most of the circular image gradients intersect. An additional post-processing step is applied to ensure that the iris center falls within the black area of the pupil.

Appearance based methods are based on the response of various image filters applied on the peri-ocular region [27]. Finally, hybrid methods [28] combine shape based methods with appearance based methods to overcome their limitations and increase the system's overall performance.

Nowadays, deep learning attained impressive results in image classification tasks, and, as expected, deceit detection was also addressed from this perspective. Several works used convolutional neural networks (CNNs) to spot and recognize micro-expressions—one of the most reliable sources of deceit detection. In [29], the authors trained a CNN on the frames from the start of the video sequence and on the onset, apex and offset frames. The convolutional layers of the trained network are connected to a long-short-term-memory recurrent neural network, which is capable of spotting micro-expressions. In [30], the authors detect micro-expressions using a CNN trained on images differences. In a post processing stage, the predictions of the CNN are analyzed in order to find the micro-expression intervals and to eliminate false positives.

The work [31] presents an automatic deception detection system that uses multi-modal information (video, audio and text). The video analysis module fuses the score of classifiers trained on low level video features and high-level micro-expressions to spot micro-expressions.

This study aims at developing computer vision algorithms to automatically analyze eye movements and to compute several oculomotor metrics which showed great promise in detecting deceitful behavior. We propose a fast iris center tracking algorithm which combines geometrical and appearance based methods. We also correlate the position of the iris center with the eye corners and eyelid apexes in order to roughly estimate the gaze direction. Based on these features, we compute several ocular cues which have been proposed as cues to deceit detection in the scientific literature: the blink rate, the gaze direction and eye movement AUs' recognition.

This work will highlight the following contributions:

- The development of a blink detection system based on the combination of two eye state classifiers: the first classifier estimates the eye state (open or closed) based on the eye's aspect ratio, while the second classifier is a convolutional neural network which learns the required filters needed to infer the eye state.
- The estimation of the gaze based on the quantification of the angle between the iris and the eye center. The iris centers are detected using a shape based method which employs only image derivatives and facial proportions. The iris candidates are only selected within the internal eye contour detected by a publicly available library.
- The definition of a novel metric, the normalized blink rate deviation which is able to capture the differences between the blink rate in cases of truthful and deceitful behavior. It computes

the absolute normalized difference between a reference blink rate and the blink rate of the new session. Next, this difference is normalized with the reference blink rate in order to achieve inter-subject variability.

The remainder of this manuscript is structured as follows: in Section 2, we present in detail the proposed solution and, in Section 3, we report the experimental results we performed. Section 4 presents a discussion of the proposed system and its possible applications and, finally, Section 5 concludes this work.

2. Eye Movement Analysis System: Design and Implementation

The outline of the proposed solution is depicted in Figure 1.

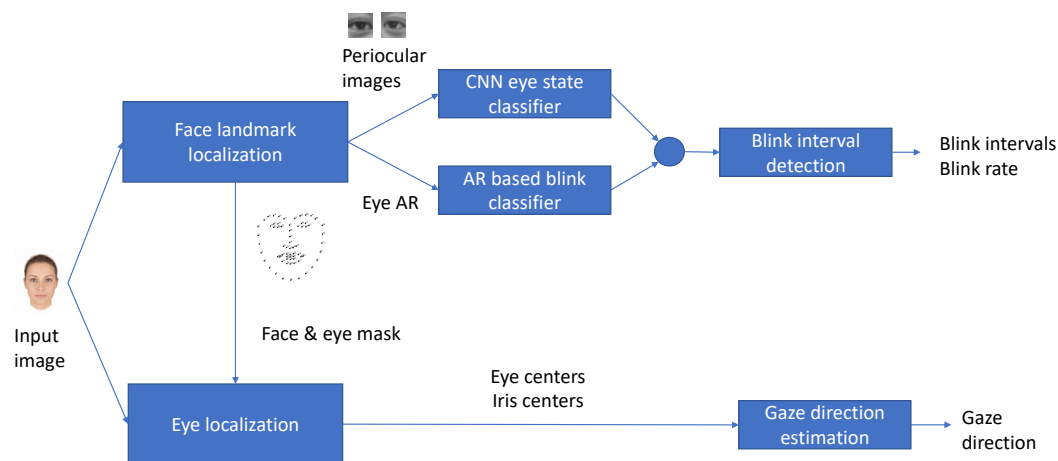


Figure 1. Solution outline. The proposed system analyses two eye features which are accepted as sources as deceit detection: blinks and gaze direction. The blink detection module combines the result of two classifiers to determine the blink intervals. The gaze analysis module determines the position of the iris centers using a shape based method and estimates the gaze direction using the distance and orientation of the eye and iris center.

The proposed eye analysis method uses a publicly available face and face landmark detection library (dlib) [32]; it detects the face area and 68 fiducial points on the face, including the eye corners and two points on each eyelid. Is the italic necessary?

These landmarks are used as the starting point of the eye movement detection and blink detection modules.

Based on the eye contour landmarks, the eye aspect ratio is computed and used as a cue for the the eye state. If the eye has a smaller aspect ratio, it is more probable that it is in the closed eye state. In addition, these landmarks are used to crop a square peri-ocular image centered in the eye center, which is fed to a convolutional neural network which detects the eye state. The response of these two classifiers is combined into a weighted average to get the final prediction on the eye state.

The eye AU type recognition module computes the iris centers and based on their relative positions to the eye corners decide on the AU.

2.1. Blink Detection

We propose a simple yet robust, appearance based algorithm for blink detection, which combines the response of two eye state classifiers: the first classifier uses the detected fiducial points in order to estimate the eye state, while the latter is a convolutional neural network (CNN) which operates on periocular image regions to detect blinks.

The first classifier analyses the eye aspect ratio ($ar = \frac{eh}{ew}$); the width and height of the eyes are computed based on the landmarks provided by the *dlib* face analysis toolkit. The width is determined

as the Euclidian distance between the inner and outer eye corner, while the height is determined as the Euclidian distance between the upper and lower eyelid apexes (Figure 2). As the face analysis framework does not compute the eyelid apexes, but two points on each eyelid, we approximate them through the interpolation between these two points. The aspect ratios extracted from each frame are stored into an array *AR*.

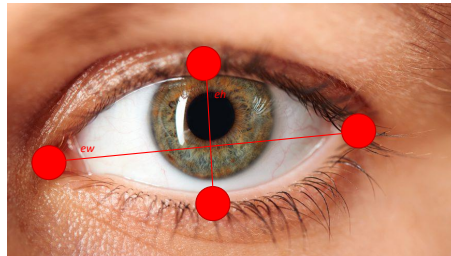


Figure 2. Eye aspect ratio. The aspect ratio $ar = \frac{eh}{ew}$ of the eyes is used to decide whether the eye is in the open or closed state.

In the frames where the eyes are closed, the aspect ratio of the eye *ar* is expected to decrease below its average value thought the video sequence. However, in our experiments, we observed that in cases of low resolution, degraded images or in the presence of occlusions (eyeglasses or hair), the fiducial points are not precisely detected, so this metric is not always reliable.

Therefore, to address this issue, we decided to combine the response of this simple classifier with the predictions of a convolutional neural network (CNN). CNNs achieved impressive results in several image recognition tasks; as opposed to classical machine learning algorithms, which require the definition and extraction of the training features, CNNs also learn optimal image filters required to solve the classification problem. To achieve time efficiency, we propose a light-weight CNN inspired from the Mobilenet architecture [33]. The key feature of Mobilenet is the replacement of classical convolutional layers with depth-wise convolutional layers, which factorize the convolutions into a depth-wise convolution followed by a point-wise convolution (1×1 convolution). These filters allow building simpler, lighter models which can be run efficiently on computational platforms with low resources, such as embedded systems and mobile devices.

The topology of the proposed network is reported in Table 1.

Table 1. Structure of the eye state detection network.

Layer	Filter Size	Stride	Number of Filters
Convolutional	3×3	2	32
Depth-wise convolutional	3×3	1	64
Depth-wise convolutional	3×3	2	128
Depth-wise convolutional	3×3	1	1024

The network has only four convolutional layers. The input layer consists of 24×24 gray-scale images and it is followed by a classical 3×3 convolutional layer. Next, three 3×3 depth-wise convolutional layers follow. A dropout layer with a dropout keep probability of 0.99 is added before the last convolutional layer which is responsible with the eye state classification. Finally, the output layer is a softmax layer.

The training data for the network consists of images from Closed Eyes In The Wild (CEW) database [34]. The dataset comprises 24×24 resolution images from 2423 subjects; 1192 subjects have the eyes closed, while the remaining 1231 have opened eyes. Some examples of images used to train the neural network are depicted in Figure 3. To ensure that the training set is representative enough and to avoid over-fitting, the training images underwent several distortions: contrast and brightness enhancement, random rotation and random crops.

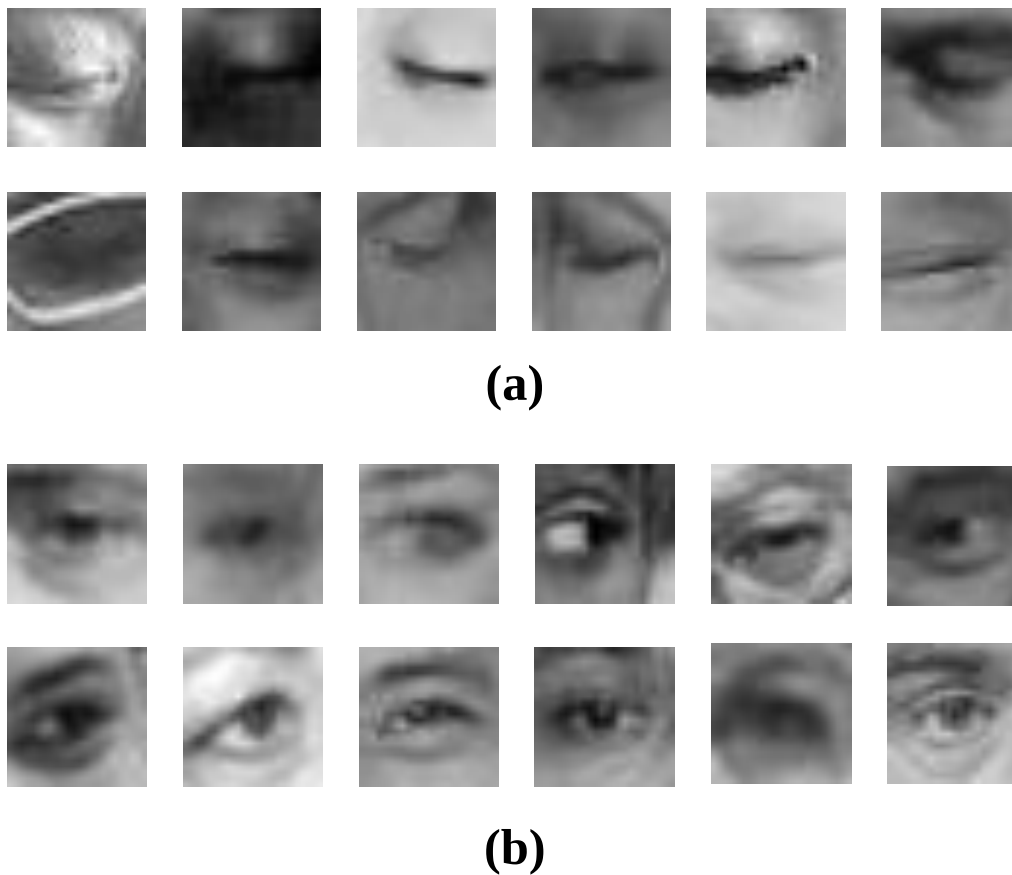


Figure 3. Examples of eye samples used to train the eye state detection convolutional neural network. (a) closed eye samples; (b) opened eye samples.

The network was trained using the softmax cross entropy loss function, RMSProp optimizer and asynchronous gradient descent, as described in [33].

The problem now is to merge the response of the two classifiers to obtain the eye state: for each frame index, we combine the two predictions into a weighted average:

$$r(t) = \alpha * p_{cnn}(t) + (1 - \alpha) * (1 - p_{ar}(t)/M), \quad (1)$$

where $r(t)$ is the response of the combined classifiers at frame t , $p_{cnn}(t)$ is the probability of the eye being closed at frame t as predicted by the CNN, $p_{ar}(t)$ is the eye aspect ratio at frame t and M is the maximum value from the array p_{ar} .

The result of combining the estimation of these two classifiers is depicted in Figure 4. In this figure, the aspect ratio is normalized as described in the above equation: it is divided by the maximum value of the p_{ar} so that its maximum value becomes 1.0 and then inverted by subtracting its value from 1.0. In this way, blinks correspond to higher values in the feature vector.

The value of the weight α was determined heuristically through trial and error experiments. All our experiments were performed using the same value for $\alpha = 0.75$, independently of the test database. We observed that the CNN gave predictions around 0.5–0.6 in case of false positives and very strong predictions (>0.97) for true positives. The aspect ratio based classifier failed to identify the closed eye state in degraded cases. We concluded that the CNN classifier should have a slightly higher weight, as in the majority of the cases it recognized with high probability the eye state. On the other hand, in cases of false positives, when combined with the other classifier, the overall classification was better as it compensated the “errors” made by the CNN.

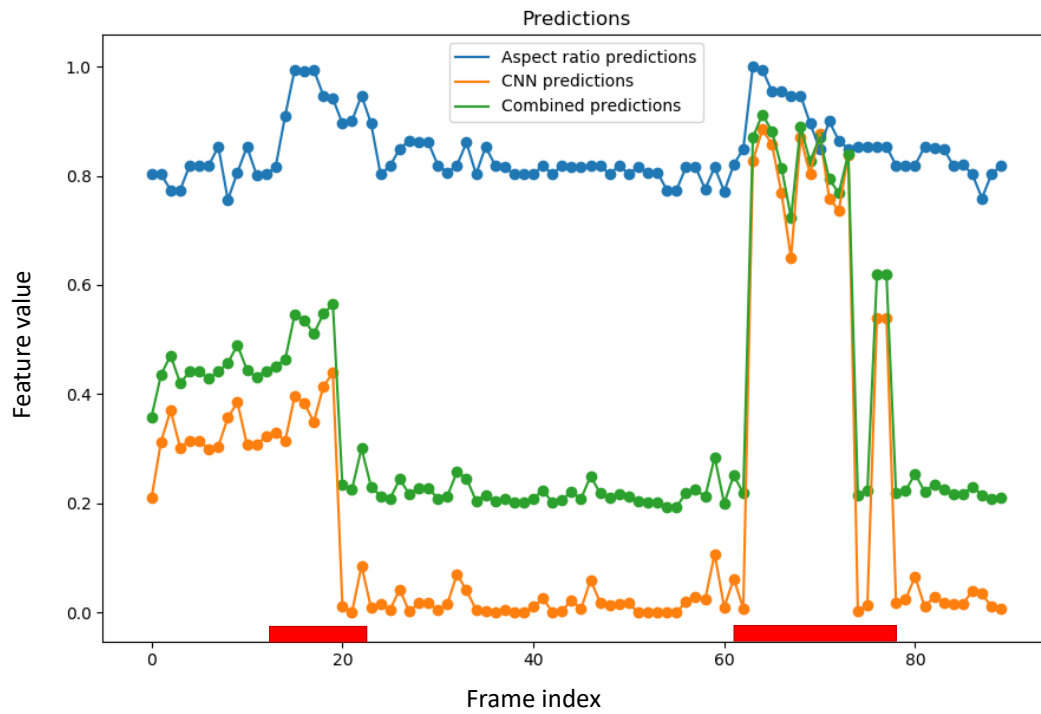


Figure 4. The eye state is estimated by combining the predictions of two classifiers based on the aspect ratio of the eye and on the response of a CNN, respectively. The ground truth position of the blinks is marked with a red bar at the bottom of the image.

The blinks (the intervals when the eyes were closed) should correspond to local maxima in the response vector R . Using a sliding window of size w , we iterate through these predictions to find the local maxima, as described in Algorithm 1. An item k from the prediction sequence is considered a local maximum if it is the largest element from the interval $[w - k, w + k]$ and its value is larger than the average τ of the elements from this interval by a threshold TH .

Algorithm 1: Blink analysis.

Input:

AR —eye aspect ratio evolution over time

w —window size

Output:

Blink rate. Blink duration.

initialize blink apexes: $BA \leftarrow \{\}$;

$sz \leftarrow \text{length}(AR)$;

while not at end of sequence **do**

$\tau \leftarrow$ average value of samples from AR in interval $[k - w, k + w]$;

if $AR[k] \leq \tau - TH$ **and** $AR[k]$ is local maximum in $[k - w, k + w]$ **then**

 insert k into BA ;

end

end

Finally, we apply a post processing step in order to avoid false positives: for each local maximum, we also extract the time interval in which the eyes were closed and we check that this interval is larger or equal to the minimum duration of an eye blink.

The blink rate (BR) is computed as the number of detected blinks (i.e., local minima from the *AR* sequence) per minute. This metric has implications in various important applications: deceit detection, fatigue detection, understanding reading and learning patterns, just to name a few.

2.2. Gaze Direction

Another deception metric computed by the system is the gaze direction. The process of determining this value implies two main steps: iris center localization and gaze direction estimation. The iris centers are detected using a shape-based eye detection method. The gaze direction is determined by analyzing the geometrical relationship between the iris center and the eye corners.

2.2.1. Iris Center Localization

Iris centers are detected using a method similar to [23]: Fast Radial Symmetry transform (FRST) [35] and anthropometric constraints are employed to localize them. FRST is a circular feature detector which uses image derivatives to determine the weight that each pixel has to the symmetry of the neighboring pixels, by accumulating the orientation and magnitude contributions in the direction of the gradient. For each image pixel p , a positively affected pixel p_+ and a negatively affected pixel p_- are computed (Figure 5); the positively affected pixel is defined as the pixel the gradient is pointing to at a distance r from p , and the negatively affected pixel is the pixel the gradient is pointed away from at a distance r from p (Figure 5).

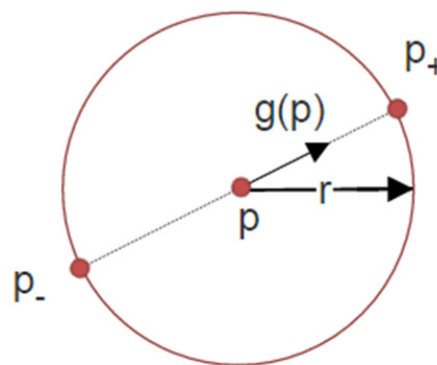


Figure 5. Positive and negatively affected pixels: the positively affected pixel (p_+) is the pixels the gradient is pointing to at r distance away from p , while the negatively affected pixel (p_-) is the pixel the gradient is pointing away from at distance r .

The transform can be adapted to search only for dark or bright regions of symmetry: dark regions can be found by considering only the negatively affected pixels, while bright regions are found by considering only positively affected pixels.

One of the main issues of the method [23] is that in cases of degraded images or light eye colors, the interior parts of the eyebrows give stronger symmetry responses than the actual iris centers, and thus lead to an inaccurate localization. In order to address this problem, we modified the method such that only the iris candidates located within the interior contour of the eye detected by the *dlib* library are selected.

First, the FRST transform of the input image is computed; the search radii are estimated based on facial proportions: the eye width is equal to approximately one fifth of the human face, while the ratio between the iris radius and the eye width is 0.42 [23].

To determine the iris candidates, the area of the FRST image within the internal contour of the eye is analyzed and the first three local minima are retained. In order to ensure that the detected minima don't correspond to the same circular object, after a minimum is detected, a circular area around it is masked so that it will be ignored when searching for the next minimum.

All the possible iris pairs are generated, and the pair with the best score is selected as the problem's solution. The score of a pair is computed as the weighted average of the pixels values from the symmetry transform image S located at the coordinates of the left (cl_x, cl_y) and right iris candidates (cr_x, cr_y) , respectively:

$$score_{pair} = \frac{S_{(cl_x, cl_y)} + S_{(cr_x, cr_y)}}{2}. \quad (2)$$

After this coarse approximation of the iris center, to ensure that the estimation is located within the black pupil area, a small neighborhood equal to half the iris radius is traversed and the center of set as the darkest pixel within that area.

2.2.2. Gaze Direction Estimation

NLP practitioners consider that the direction of the eyes can indicate whether a person is constructing or remembering information. In addition, in the FACS methodology, the movements of the eyes are encoded into the following AUs: AU61 and AU62—eyes positioned to the left and right respectively, and AU63 and AU64, eyes up or eyes down.

The proposed system recognizes the four AUs which describe the eye movements: for each frame of the video sequence, we compute the angle between the center of the eyes (computed as the centroid of the inner eye contour detected by *dlib*) and the center of the iris:

$$\theta = \tan^{-1} \frac{ec.y - ic.y}{ec.x - ic.x}, \quad (3)$$

where ec and ic are the coordinates of the eye and iris center respectively, and θ is the angle between these two points.

The next step is the consists in the quantization of these angles to determine the eye AU, as illustrated in Figure 6.

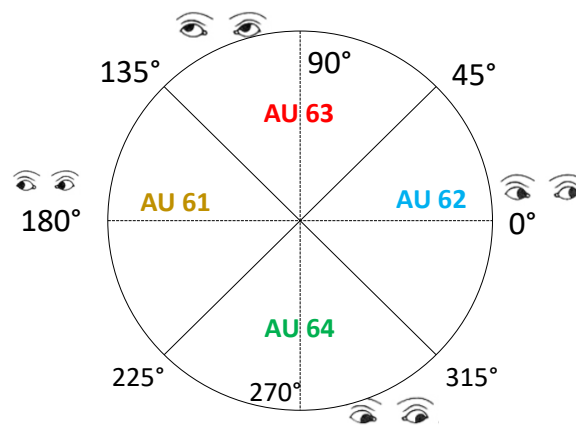


Figure 6. Quantization of the angle between the iris center and the eye center into eye movement action units.

We defined some simple rules based on which we recognize the eye movement action units. For AU61, eye movement to the left, the distance between the iris center and the eye center must be larger than TH_h and the angle (in degrees) between these two points must be between $[135^\circ, 225^\circ]$. For AU62, eye movement to the right, the angle between the iris and the eye center must lie in one of the intervals: $[0^\circ, 45^\circ]$ or $[315^\circ, 360^\circ]$. In addition, the distance between these two points must be larger than TH_h . The value of TH_v was determined heuristically and expressed in terms of facial proportions: $0.075 * ipd$, where ipd is the inter-pupillary distance.

3. Experimental Results

3.1. Blink Detection

The proposed eye blink detection algorithm was tested on three publicly available datasets.

The ZJU Eyeblink database [19] comprises 80 video (10876 frames) sequences which contain in total 255 blinks, captured at 30 frames per second (FPS) and 320×240 image resolution. The subjects were asked to blink naturally, in different experimental setups: without wearing eyeglasses: frontal and upward view, and wearing glasses, either thin rim or black frame glasses. The average duration of a video sequence is 5 s. The authors of [14] performed an independent manual annotation of the dataset and reported 272 eye blinks. This difference could be caused by the fact that in some videos, the subjects blink several times very fast and could be considered as multiple blinks.

The Eyeblink8 database was introduced in [18] and it poses more difficult challenges, as the subjects were asked to perform (natural) face movements which can influence the performance of blink detection algorithm. The database is comprised of eight video sequences (82,600) captured at a normal frame rate and 640×480 image resolution; in total, 353 blinks were reported. Each frame is annotated with the state of the eyes, which can have one of the following values: open, half-open or close.

The Silesian face database [36] is comprised of more than 1.1 million frames captured with a high speed camera at 100 FPS and 640×480 image resolution. The main purpose of the database is to provide the scientific community a benchmark of facial cues to deceit. The participants of the study were third and fourth year students (101 subjects); they were told that they will help with assessing the power of a person who claims to be a telepath; the experiment implied that they should tell the truth or lie about some geometrical shapes displayed on the computer screen. Therefore, the participants were not aware about the purpose of the study, so all the facial movements were genuine. Each frame from the dataset was carefully annotated by three trained coders with different non-verbal cues which can indicate deception: eye blinks (more specifically eye-closures), gaze aversion and micro-expressions (facial distortions).

An important problem when evaluating the performance of a blink detection system is how to compute the number of negatives (false negatives or true negatives) from a video sequence. We used the same convention as in [18]: the number of non-eye blinks is computed by dividing the number of frames with open eyes from the dataset with the average blink duration (expressed in video frames; an average blink takes 150–300 ms, so between 5–10 frames at 30 fps).

We report the following metrics: precision $Pr = \frac{TP}{TP+FP}$, recall $R = \frac{TP}{TP+FN}$, the false positive rate $FPr = \frac{FP}{FP+TN}$ and the accuracy $ACC = \frac{TP+TN}{FP+FN+TP+TN}$.

Table 2 reports the performance of the blink detection module on three state of the art datasets. As mentioned above, the original ZJU Eyeblink dataset does not contain the temporal annotations of the blink intervals, as the authors only specify the number of blinks and the average blink duration. We ran the proposed algorithm on the ZJU Eyeblink datasets and it detected 263 blinks.

Table 2. Blink detection algorithm performance.

Database	Precision	Recall	FPr	Accuracy
Eyeblink 8	94.75%	94.89%	0.38%	99.30%
Silesian	100%	96.68%	0%	99.41%

From Table 2, it can be noticed that the proposed eye blink detection algorithm attained 100% precision on the Silesian dataset; in other words, there were no false positives reported. The recall rate is also high, meaning that the proposed solution does not miss many blink intervals. A comparison of the proposed eye blink detection method with the state of the art is reported in Table 3.

Table 3. Iris center results compared to other methods.

Method	Database	Precision	Recall
[37]	Eyeblink8	94.69%	91.91%
Proposed solution	Eyeblink8	94.75%	94.89%

The majority of other works presented in the literature report their results on the ZJU Eyeblink dataset. However, the original version of this dataset is not annotated with the blink intervals. Some works tried to manually annotate the ZJU Eyeblink dataset and obtained a different number of blinks. For example, if the authors reported 255 blinks, in [14], the authors counted 272 blink intervals. It is worth mentioning that the Eyeblink8 dataset poses more challenges than the ZJU dataset as the subjects perform other facial expressions and head movements.

When evaluating on the Eyeblink8 dataset, compared to [37], we obtained a slightly higher precision value (0.06%), and the recall metric is almost 3% higher. This increase in the recall value means that the proposed solution achieves a lower number of false negatives; this could be a consequence of the fact that we use more complex classifiers (CNN) which are able to spot the closed eyes in more difficult scenarios. The precision metric is high 94.75%, but not very different from the one obtained by [37]. This indicates that we obtain a similar number of false positives as [37]. We noticed that our method detects false positives in image sequences in which the subjects perform other facial expressions in which the eyes are almost closed, such as smiling or laughing. We argue that this problem could be addressed by increasing the training set of the CNN with periocular images in which the subjects laugh or smile, as now the training set contains only images from the CEW dataset in which the participants have an almost neutral facial expression.

3.2. Iris Center Localization: Gaze Direction

The BIO-ID face database [38] is often used as a benchmark for iris center localization by the research community. It contains 1521 gray-scale images of 23 subjects captured under different illumination settings and with various eye occlusions (eyeglasses, light reflexions, eyes closed). Each image contains manual annotations of the eye positions.

In [39], the authors proposed an iris center localization evaluation metric which can be interpreted independently of the image size: the maximum distance to the ground truth eye location obtained by the worst of both eye estimators, normalized with the interpupillary distance *wec*—worst eye center approximation:

$$wec = \frac{\max(\|\tilde{C}_l - C_l\|, \|\tilde{C}_r - C_r\|)}{\|C_l - C_r\|}, \quad (4)$$

where C_l, C_r are the ground truth positions of the left and right iris center, and $\tilde{C}_l, \tilde{C}_r$ are their respective estimates.

This metric can be interpreted in the following manner: if $wec \leq 0.25$, the error is less than or equal to the distance between the eye center and the eye corners, if $wec \leq 0.10$, the localization error is less than or equal to the diameter of the iris, and, finally, if $wec \leq 0.05$, the error is less than or equal to the diameter of the pupil.

Two other metrics can be simply derived from *wec*: *bec* (best eye center approximation) and *aec* (average eye center approximation) which define the lower and the averaged error, respectively:

$$bec = \frac{\min(\|\tilde{C}_l - C_l\|, \|\tilde{C}_r - C_r\|)}{\|C_l - C_r\|}, \quad (5)$$

$$aec = \frac{\text{avg}(\|\tilde{C}_l - C_l\|, \|\tilde{C}_r - C_r\|)}{\|C_l - C_r\|}, \quad (6)$$

where *avg* and *min* are the average and minimum operators.

The Talking Face video [40] was designed as a benchmark for face analysis methods developed to analyze facial movements in natural conversational scenarios. It contains 5000 video frames corresponding to approximately three and a half minutes of conversation. Each frame from the image sequence is annotated in a semi-automatic manner with the positions of 68 facial landmarks, including the iris centers.

The results obtained by the proposed method on the BIO-ID and Talking Face databases are reported in Table 4.

Table 4. Performance of the proposed iris center localization method on the BIO-ID and Talking Facedatabases.

Dataset	Error $\leq 0.05(\%)$			Error $\leq 0.10(\%)$			Error $\leq 0.25(\%)$		
	<i>bec</i>	<i>aec</i>	<i>wec</i>	<i>bec</i>	<i>aec</i>	<i>wec</i>	<i>bec</i>	<i>aec</i>	<i>wec</i>
BIO-ID	91.47	86.92	80.96	94.26	92.79	91.47	99.92	99.70	99.19
Talking face	96.90	94.97	89.59	98.09	97.81	97.60	99.97	99.97	99.95

From Table 4, it can be noticed that, in more than 86% (BIO-ID face database) and 94% (Talking facedatabase) of the cases, the average iris center approximation falls within the iris area. The results are better on the Talking Face database; this is as expected as the image resolution is higher (720×576 vs. 384×286 in the BIO-ID database) and also the face fills a higher area in the images of the Talking Face dataset. In addition, the frames from the BIO-ID dataset contain more occlusions: eye semi-closed, high specular refraction on glasses, etc. However, both datasets were captured in unconstrained scenarios. In the Talking Face dataset, the subject was allowed to move his head freely and express his emotion.

Figures 7 and 8 illustrate the cumulative error distribution plots for the worst, average and best eye approximations on the datasets.

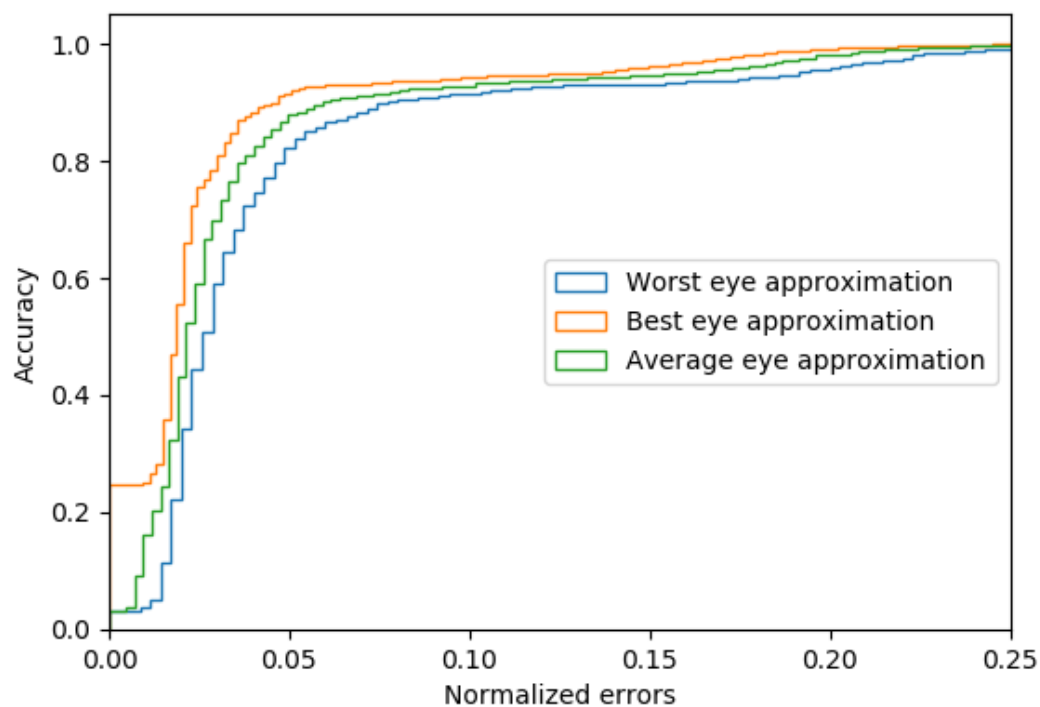


Figure 7. Cumulative error distribution on the BIO-ID face database.

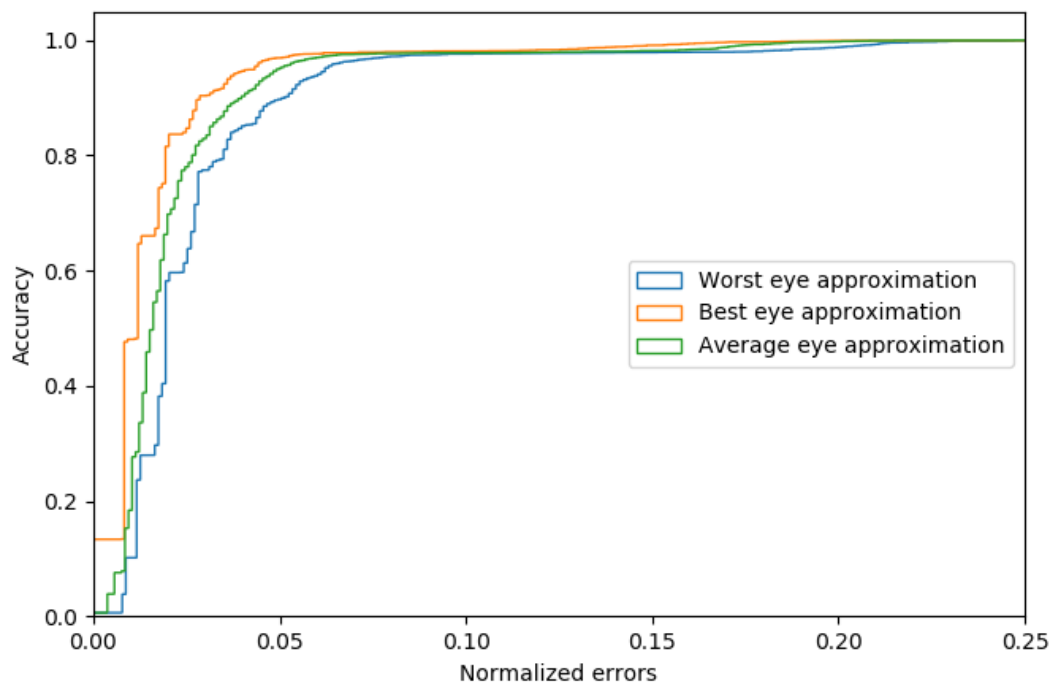


Figure 8. Cumulative error distribution on the Talking Face database.

A comparison of the proposed method with other iris center detection methods is reported in Table 5. For the methods marked with an *, the accuracy that was read from the performance curves as a numerical value was not provided in the manuscript.

Table 5. Iris center localization results compared to other methods.

Method	$wec \leq 0.05(\%)$	$wec \leq 0.10(\%)$	$wec \leq 0.25(\%)$
[26]	82.5%	93.4%	98.0%
[39]	38.0%*	78.8%*	91.8%
[28]	57%*	96%	97.1%
[23]	74.65%	79.15%	98.09%
Proposed Solution	91.47%	94.26%	99.19%

Our iris localization method is based on [23], but it uses a smaller input space by including information about the eye corners and eyelid apexes provided by the *dlib* library. The performance of the worst eye approximation (*wec*) surpasses [23] by more than 6% for the case $wec \leq 0.05$ (i.e., the iris center is detected within the pupil area). As expected, the case $wec \leq 0.25$ (the iris center approximation is within the eye area) is close to 100% (99.92%), as the search area includes only the internal eye contour. In conclusion, the proposed iris localization algorithm exceeds the other methods in all the experimental setups: within pupil localization ($wec \leq 0.05$), within iris localization ($wec \leq 0.10$) and within eye localization ($wec \leq 0.25$). Our method obtained better results mainly because the region of interest for the iris centers is restricted only to the interior of the eye. One problem of [23] was that in cases of light colored irises, the interior corners of the eyebrows gave stronger circular response than the irises; by setting a smaller search region, we filtered out these false positives.

4. Discussion

In this section, we provide a short discussion on how the proposed method could be used in deceit detection applications.

The blink rate has applications in multiple domains: driver monitoring systems, attention assessment.

The Silesian face database was developed in order to help researchers investigate how eye movements, blinks and other facial expressions could be used as cues to deceit detection. It includes annotations about the blink intervals, small eye movements and micro-tensions. A description of how the Silesian database was captured is provided in Section 3.

As some studies suggest that there is a correlation between the blink rate and the difficulty of the task performed by the subject, we searched to see if there are any differences between the blink rates based on the number of questions answered wrong by the participants from the Silesian face database. We assumed that the subjects that performed multiple mistakes found the experiment more complex.

On average, the participants answered incorrectly (made mistakes on) 0.762 from the 10 questions. The histogram of the mistakes made by the participants depicted in Figures 9 and 10 shows the average blink rate based on the number of mistakes made by the participants.

The subjects that correctly answered all the questions and those who made 1, 2 and 3 mistakes have approximately the same blink rates. The subject who got 4 out of 10 questions wrong had a significantly higher blink rate. Strangely, the subject who made the most mistakes (7 of the 10 questions were not answered correctly) had a lower blink rate. Of course this isn't necessarily statistically relevant because of the low number of subjects (one person got four questions wrong and another one got seven questions wrong). We also analyzed the correlation between the blink rate and deceitful behavior. In [36], the authors report the average blink rate per question, without taking into consideration the differences that might occur due to inter-subject differences. Based on this metric, there wasn't any clear correlation between the blink rate and the questions in which the subject lie about the shape displayed to them on the computer screen.

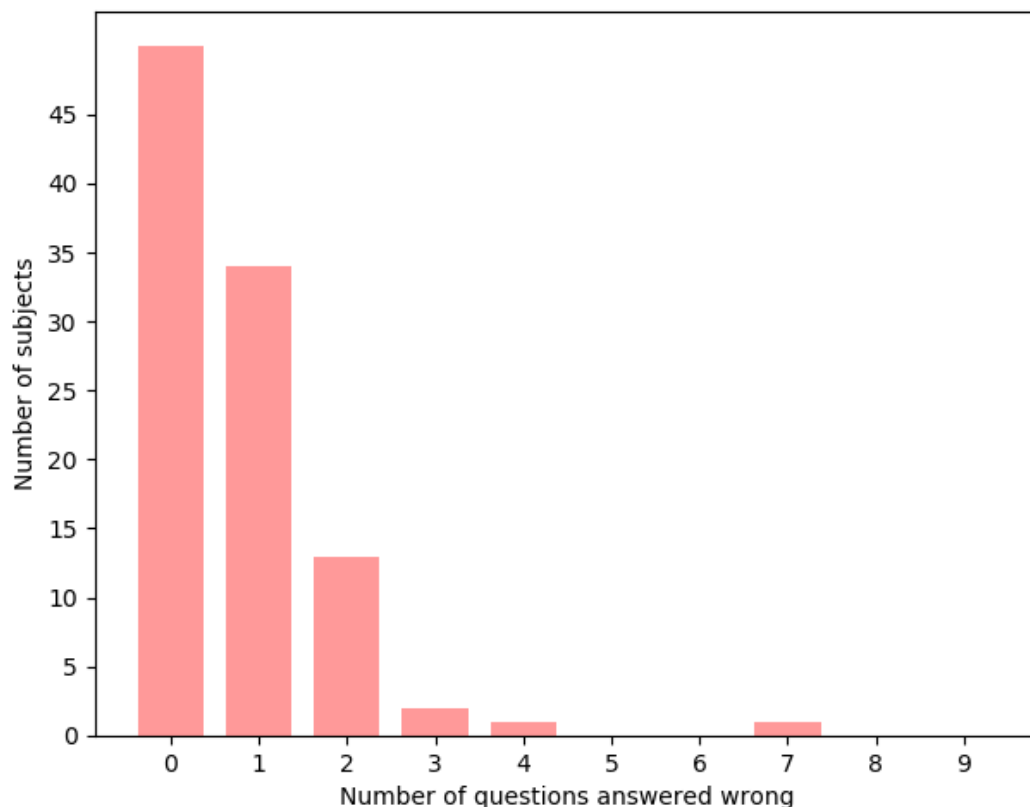


Figure 9. Histogram of the mistakes made by the participants.

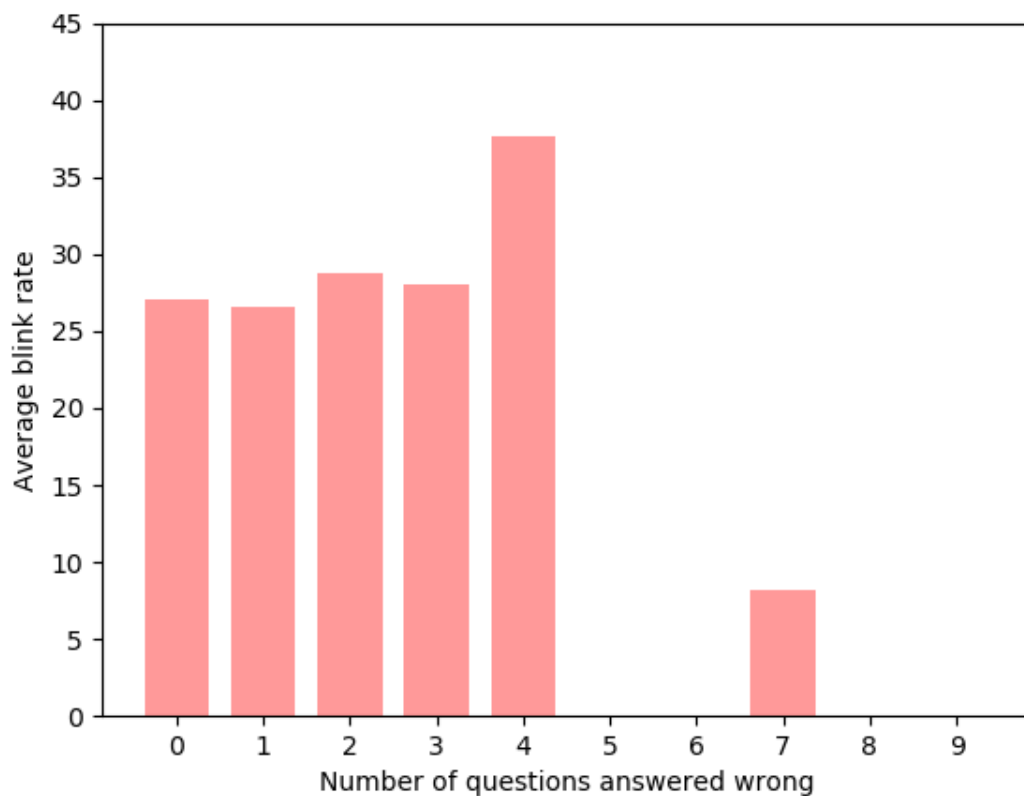


Figure 10. Average blink rate related to the number of mistakes made by the participants.

We propose another metric that takes into consideration the inter-subject variability: the normalized blink rate deviation *NBRD*. As there are less questions in which the subjects tell the truth (three truthful answers vs. seven deceptive answers), we compute for each subject a reference blink rate br_{ref} as the average of the blink rates in the first four deceptive answers. For the remaining six questions, we compute the *NBRD* metric as follows:

$$NBRD_q = \frac{abs(br_q - br_r)}{br_r}, \quad (7)$$

where br_q is the blink rate for the current question and br_r is the reference blink rate computed for the subject as the average blink rate on the four deceptive answers. This methodology is somewhat similar to the Control Question Technique [2], as for each new question we analyze the absolute blink rate difference to a set of “control” questions. Are the italics necessary?

The next step is to apply a simple classifier (a Decision Stump in our case) and to see if the deceptive answers could be differentiated from the truthful ones. We randomly selected $n_{tf} = 49$ subjects to train the decision stump and we kept the other $n_t = 52$ to test the accuracy of the decision stump. For each subject, we have three truthful questions, three deceptive questions (the other four deceptive questions are kept as a normalization reference factor), so the test data is balanced.

The performance of the simple classifier is reported in Table 6 and the corresponding confusion matrix is reported in Table 7.

Table 6. Deceitful behavior detection based on the blink rate.

Classifier	Precision	Recall	Accuracy	F1-Score
Decision stump	96.75%	95.51%	96.15%	96.12%

Table 7. Confusion matrix for deceitful behavior detection based on blink rates.

Ground truth/Detected	Deception	Truth
Deception	151	5
Truth	7	149

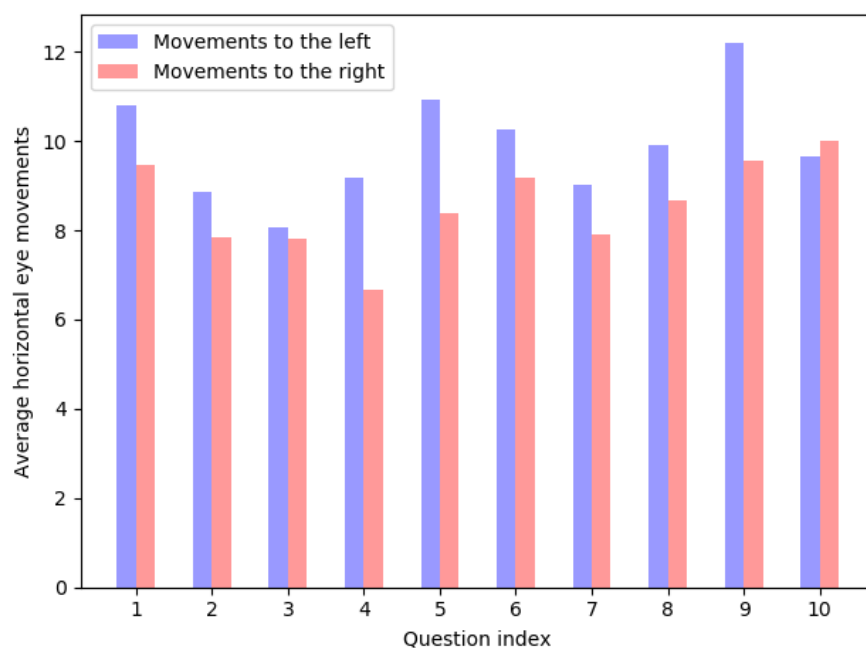
Therefore, from Table 6, we can conclude that a simple classifier was able to differentiate between the truthful and the deceitful questions based on the proposed *NBRD* metric.

Regarding the gaze direction in the deceit detection, to our knowledge, there isn't any database annotated with the gaze direction in the context of a deceit detection experiment. The Silesian face database provides annotation with the subtle movements of the gaze (saccades); the following dictionary of eye movement is defined: EyeLeft, EyeRight, EyeUp, EyeDown, Neutral, EyeLeftUp, EyeLeftDown, EyeRightUp, EyeRightDown. However, these movements are short, low-amplitude, ballistic eye movements, and differ a lot from "macro" eye movements.

In the experiment, the subjects are asked to respond to the questions of a person that they believe is a telepath, according to some instructions displayed on the screen. Therefore, the experimental setup is more controlled and the subjects don't need to access their memory (they need to provide a predefined answer), so it is normal that "macro" movements do not occur. The participants were instructed to tell the truth for questions 1, 2 and 9 and to lie for all the other questions.

However, we analysed the provided saccadic data in order to determine if there is any connection between direction of the saccades and deceitful behavior. However, as opposed to [36], we don't analyze the number of gaze aversions, but we split the gaze aversion into four classes corresponding to the eye movements defined in the FACS methodology: eyes up, eyes down, eyes left and eyes right. NLP theories claim that these movements could indicate deceit.

Figure 11 illustrates the average lateral eye movements (left—annotations EyeLeft, EyeLeftUp, EyeLeftDown—vs. right—annotations EyeRight, EyeRightUp, EyeRightDown), while Figure 12 illustrates the vertical eye movements (up and down, annotations EyeUp vs. EyeDown, respectively).

**Figure 11.** Average lateral saccadic gaze shifts on the Silesian face database. The subjects told the truth on questions 1, 2, and 9 and lied for the other ones.

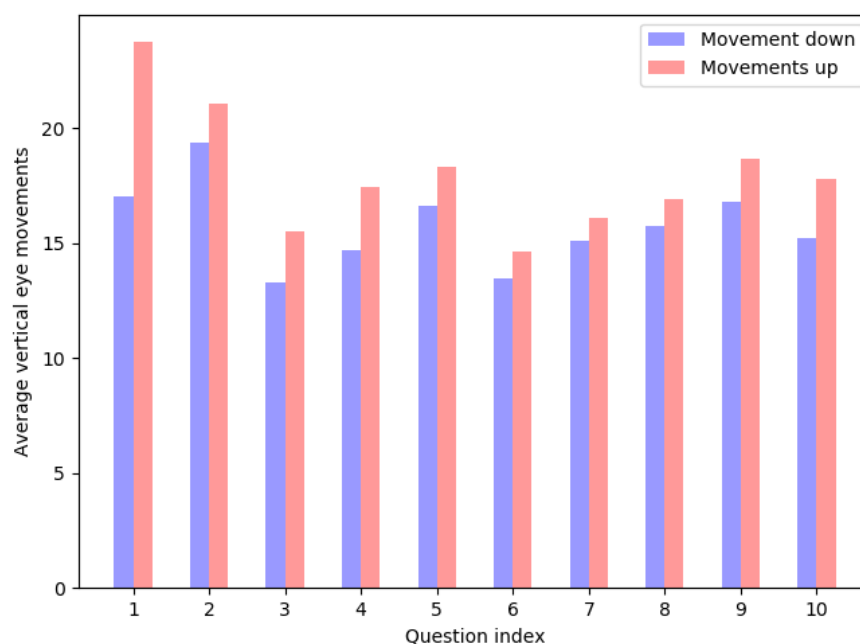


Figure 12. Average vertical saccadic gaze shifts on the Silesian face database. The subjects told the truth on questions 1, 2, and 9 and lied for the other ones.

There doesn't seem to be any distinguishable pattern in saccadic eye movements which could indicate deceit. We also applied the proposed deceit detection metric NBRD and a simple decision stump to try to detect the deceitful detection based on eye movements. The results were not so satisfying: we obtained classification rates worse than average on both left eye movements and right eye movements. However, it is worth mentioning that the saccades from this database are not necessarily non-visual saccades (the type of saccades that has been correlated to deceit), but visual saccades (the student needs to read a predefined answer from a computer screen).

5. Conclusions

In this manuscript, we presented an automatic facial analysis system that is able to extract various features about the eyes: the iris centers, the approximative gaze direction, the blink intervals and the blink rate.

The iris centers are extracted using a shape based method which exploits the circular (darker) symmetry of the iris area and facial proportions to extract to locate the irises. Next, the relative orientation angle between the eye center and the detected iris center is analyzed in order to determine the gaze direction. This metric can be used as a cue to deception: the interlocutor not making eye contact and often shifting his gaze point could indicate that he feels uncomfortable or has something to hide. In addition, some theories in the field of NLP suggests that the gaze direction indicates if a person remembers or imagines facts.

The proposed system also includes a blink detection algorithm that combines the response of two classifiers to detect blink intervals. The first classifier relies on the eye's aspect ratio (eye height divided by eye width), while the latter is a light-weight CNN that detects the eye state from peri-ocular images. The blink rate is extracted as the number of detected blinks per minute.

The proposed solution was evaluated on multiple publicly available datasets. On the iris center detection task, the proposed method surpasses other state-of-the-art works. Although this method uses the same image feature as [23] to find the circular iris area, its performance was boosted with more than 6% by selecting the input search space based on the position of the facial landmarks extracted with the *dlib* library.

We also proposed a new deceit detection metric—NBRD—the normalized blink rate deviation, which defined the absolute difference between the blink rate from new situations normalized with the subject's average blink rate. Based on this metric, a simple decision stump classifier was able to differentiate between the truthful and the deceitful questions with an accuracy of 96%.

As a future work, we plan to capture a database intended for deceit detection; our main goal is to let the subjects interact and talk freely in an interrogation-like scenario. For example, the subjects will be asked to randomly select a note which contains a question and an indication to whether they should answer that question truthfully or not. Next, they will read the question out loud and discuss it with an interviewer; the interviewer—who is not aware if the subject is lying or not—will engage in an active conversation with the participant. In addition, we intend to develop robust motion descriptors based on optical flow that could capture and detect the saccadic eye movements. We plan to detect the saccadic eye movements using a high speed camera by analyzing the velocity of the detected eye movements.

Author Contributions: Conceptualization, R.D. and D.B.; Software, D.B.; Validation, R.I., D.B. and R.D.; Formal Analysis, R.D.; Resources, R.D.; Writing—Original Draft Preparation, B.D.

Funding: This research was funded by the Romanian National Authority for Scientific Research, CNDI-UEFISCDI, Grant No. PN-III-P1-1.1-TE2016-0440—DEEPESENSE.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACC	Accuracy
CNN	Convolutional neural network
FACS	Facial action coding system
FP	False positives
FPS	Frames per second
FRST	Fast radial symmetry transform
NBRD	normalized blink rate deviation
NLP	Neuro-linguistic programming
Pr	Precision
R	Recall
TP	True positives

References

- Ekman, P. Lying and nonverbal behavior: Theoretical issues and new findings. *J. Nonverbal Behav.* **1988**, *12*, 163–175.
- Ben-Shakhar, G. Current research and potential applications of the concealed information test: An overview. *Front. Psychol.* **2012**, *3*, 342.
- Lykken, D.T. The GSR in the detection of guilt. *J. Appl. Psychol.* **1959**, *43*, 385.
- Ekman, P. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage (Revised Edition)*; WW Norton & Company: New York, NY, USA, 2009.
- Bandler, R.; Grinder, J. *The Structure of Magic*; Science and Behavior Books: Palo Alto, CA, USA, 1975; Volume 1.
- Wiseman, R.; Watt, C.; TenBrinke, L.; Porter, S.; Couper, S.L.; Rankin, C. The eyes don't have it: Lie detection and neuro-linguistic programming. *PLoS ONE* **2012**, *7*, e40259.
- Witkowski, T. Thirty-five years of research on Neuro-Linguistic Programming. NLP research data base. State of the art or pseudoscientific decoration? *Pol. Psychol. Bull.* **2010**, *41*, 58–66.
- Ekman, P.; Friesen, W.V.; Hager, J.C. *Facial Action Coding System. Manual and Investigator's Guide*; Nexus: Salt Lake City, UT, USA; 2002.
- Stern, J.A.; Walrath, L.C.; Goldstein, R. The endogenous eyeblink. *Psychophysiology* **1984**, *21*, 22–33.

10. Schiffman, H.R. *Sensation and Perception: An Integrated Approach*; John Wiley & Sons: Oxford, UK, 1990.
11. Gamer, M.; Pertzov, Y. Detecting concealed knowledge from ocular responses. In *Detecting Concealed Information and Deception*; Academic Press: Cambridge, MA, USA, 2018; p. 169.
12. Marchak, F.M. Detecting false intent using eye blink measures. *Front. Psychol.* **2013**, *4*, 736.
13. Peth, J.; Kim, J.S.; Gamer, M. Fixations and eye-blinks allow for detecting concealed crime related memories. *Int. J. Psychophysiol.* **2013**, *88*, 96–103.
14. Danisman, T.; Bilasco, I.M.; Djeraba, C.; Ihaddadene, N. Drowsy driver detection system using eye blink patterns. In Proceedings of the International Conference on Machine and Web Intelligence, Algiers, Algeria, 3–5 October 2010; pp. 230–233.
15. Jiménez-Pinto, J.; Torres-Torriti, M. Optical flow and driver's kinematics analysis for state of alert sensing. *Sensors* **2013**, *13*, 4225–4257.
16. Królak, A.; Strumillo, P. Eye-blink detection system for human–computer interaction. *Univ. Access Inf. Soc.* **2012**, *11*, 409–419.
17. Oh, J.; Jeong, S.Y.; Jeong, J. The timing and temporal patterns of eye blinking are dynamically modulated by attention. *Hum. Mov. Sci.* **2012**, *31*, 1353–1365.
18. Drutarovsky, T.; Fogelton, A. Eye blink detection using variance of motion vectors. In *Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014*; Springer: Berlin, Germany, 2014; pp. 436–448.
19. Pan, G.; Sun, L.; Wu, Z.; Lao, S. Eyeblink-based anti-spoofing in face recognition from a generic webcam. In Proceedings of the 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 14–21 October 2007.
20. Sukno, F.M.; Pavani, S.K.; Butakoff, C.; Frangi, A.F. Automatic assessment of eye blinking patterns through statistical shape models. In *International Conference on Computer Vision Systems, Liège, Belgium, 13–15 October 2009*; Springer: Berlin, Germany, 2009; pp. 33–42.
21. Lalonde, M.; Byrns, D.; Gagnon, L.; Teasdale, N.; Laurendeau, D. Real-time eye blink detection with GPU-based SIFT tracking. In Proceedings of the Fourth Canadian Conference on Computer and Robot Vision (CRV'07), Montreal, QC, Canada, 28–30 May 2007; pp. 481–487.
22. Hansen, D.; Ji, Q. In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 478–500.
23. Borza, D.; Darabant, A.S.; Danescu, R. Real-time detection and measurement of eye features from color images. *Sensors* **2016**, *16*, 1105.
24. Daugman, J. How iris recognition works. In *The Essential Guide to Image Processing*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 715–739.
25. Borza, D.; Danescu, R. Eye shape and corners detection in periocular images using particle filters. In Proceedings of the 12th International Conference Signal-Image Technology & Internet-Based Systems (SITIS), Naples, Italy, 28 November–1 December 2016; pp. 15–22.
26. Timm, F.; Barth, E. Accurate eye centre localisation by means of gradients. *Visapp* **2011**, *11*, 125–130.
27. Radu, P.; Ferryman, J.; Wild, P. A robust sclera segmentation algorithm. In Proceedings of the 7th International Conference on Biometrics Theory, Applications and Systems (BTAS), Virginia, VA, USA, 8–11 September 2015; pp. 1–6.
28. Cristinacce, D.; Cootes, T.F.; Scott, I.M. A multi-stage approach to facial feature detection. *BMVC* **2004**, *1*, 277–286.
29. Breuer, R.; Kimmel, R. A deep learning perspective on the origin of facial expressions. *arXiv* **2017**, arXiv:1705.01842, preprint.
30. Imai, F.H.; Trémeau, A.; Braz, J. In *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018)—Volume 5: VISAPP, Funchal, Madeira, Portugal, 27–29 January 2018*; SciTePress: Setubal, Portugal, 2018.
31. Wu, Z.; Singh, B.; Davis, L.S.; Subrahmanian, V. Deception detection in videos. *arXiv* **2017**, arXiv:1712.04415, preprint.
32. King, D.E. Dlib-ml: A machine learning toolkit. *J. Mach. Learn. Res.* **2009**, *10*, 1755–1758.
33. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861, preprint.

34. Song, F.; Tan, X.; Liu, X.; Chen, S. Eyes closeness detection from still images with multi-scale histograms of principal oriented gradients. *Pattern Recognit.* **2014**, *47*, 2825–2838.
35. Loy, G.; Zelinsky, A. Fast radial symmetry for detecting points of interest. *IEEE Trans. Pattern Anal. Mach. Intell.* **2003**, *25*, 959–973.
36. Radlak, K.; Bozek, M.; Smolka, B. Silesian Deception Database: Presentation and Analysis. In Proceedings of the ACM on Workshop on Multimodal Deception Detection, Seattle, WA, USA, 9–13 November 2015; pp. 29–35.
37. Fogelton, A.; Benesova, W. Eye blink detection based on motion vectors analysis. *Comput. Vis. Image Underst.* **2016**, *148*, 23–33.
38. Jesorsky, O.; Kirchberg, K.J.; Frischholz, R.W. Robust face detection using the hausdorff distance. In *Proceedings of the International Conference on Audio-and Video-Based Biometric Person Authentication, Halmstad, Sweden, 6–8 June 2001*; Springer: Berlin, Germany 2001; pp. 90–95.
39. Bigun, J.; Smeraldi, F. *Proceedings of the Audio-and Video-Based Biometric Person Authentication: Third International Conference, AVBPA 2001, Halmstad, Sweden, 6–8 June, 2001*; Springer: Berlin, Germany, 2003; Volume 2091.
40. Cootes, T. Talking Face Video. Available online: http://www-prima.inrialpes.fr/FGnet/data/01-TalkingFace/talking_face.html (accessed on 16 October 2018).



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).