2016 23rd International Conference on Pattern Recognition (ICPR)
Cancún Center, Cancún, México, December 4-8, 2016

Learning-by-Synthesis for Accurate Eye Detection

Chao Gou[1,3], Yue Wu[2], Kang Wang[2], Fei-Yue Wang[1,3], and Qiang Ji[2]

[1]Institute of Automation, Chinese Academy of Sciences, Beijing, China
[2]ECSE Department, Rensselaer Polytechnic Institute, Troy, USA
[3]Qingdao Academy of Intelligent Industries, Qingdao, China
{gouchao2012,feiyue.wang}@ia.ac.cn {wuy9, wangk10, jiq}@rpi.edu

***Abstract*—Cascade regression framework has been successfully applied to facial landmark detection and achieves state-of-the-art performance recently. It requires large number of facial images with labeled landmarks for training regression models. We propose to use cascade regression framework to detect eye center by capturing its contextual and shape information of other related eye landmarks. While for eye detection, it is time-consuming to collect large scale training data and it also can be unreliable for accurate manual annotation of eye related landmarks. In addition, it is difficult to collect enough training data to cover various illuminations, subjects with different head poses and gaze directions. To tackle this problem, we propose to learn cascade regression models from synthetic photorealistic data. In our proposed approach, eye region is coarsely localized by a facial landmark detection method first. Then we learn the cascade regression models iteratively to predict the eye shape updates based on local appearance and shape features. Experimental results on benchmark databases such as BioID and GI4E show that our proposed cascade regression models learned from synthetic data can accurately localize the eye center. Comparisons with existing methods also demonstrates our proposed framework can achieve preferable performance against state-of-the-art methods.**

I. Introduction

Human eyes play an important role in our everyday life for communication, interaction and other daily routines. Eye detection aims to estimate the pupil location in a image. It is becoming an increasingly important research topic due to its various applications, including iris recognition, eye gaze estimation and human-computer interaction. Learning-based methods is a promising solution for eye detection. However, learning-based methods need large number of annotated training data to cover various appearance of eyes. Although much work has been done for eye detection, they still are challenging tasks due to illumination, head pose and various gaze directions.

In this paper, we propose an effective coarse-to-fine framework for eye detection, based on learning cascade regression models from synthetic eye images. Our main contributions are summarized as follows: (I) integrate the cascade regression framework for accurate pupil detection; (II) incorporate local appearance features of pupil, shape and contextual information of eyelids as well as eye corners for eye detection; (III) learn from pure synthetic data.

The remainder of this paper is arranged as follows. Related works on eye detection are reviewed in section II. Our proposed approach is described in section III. Section IV reports the experimental results with discussions. We draw conclusions in section V.

II. Related Work

Eye detection has been studied for decades and numerous methods have been proposed. In this section, we focus on reviewing most of the recent works. A detailed review of earlier techniques devoted to this topic can be found in [1], [2].

Valenti and Gevers [3] use curvature of isophotes in the intensity image and design a voting-based method for pupil localization. SIFT-based features are extracted in each candidate pupils followed by binary classification to eliminate the false centers. Timm and Barth [4] propose an objective function based on intensity gradients and squared dot products. It is maximised for iris centers of circular regions. By using the isophote and gradient features of pupils, Zhang *et al.* localize the eye center through two response map from two modalities operations on face images. These shape-based methods can fail when pupils are partly occluded or subject with head poses. Araujo *et al.* [5] propose an inner product detector for eye localization on the basis of correlation filters. This appearance-based method is robust to small variance of the desired eye images with a low computational cost. But it is limited to eyes with multiple appearance in the wild. In [6], the authors use Support Vector Regressor (SVR) to estimate the distance of patch center to the pupil center using the extracted HOG features. Zhou *et al.* [7] employ a coarse-to-fine strategy for eye center detection, on the basis of Supervised Decent Method (SDM) [8]. They jointly use two eye related 14 points of multi-scale feature to capture the contextual information for eye center detection on given eye regions. In [9], Wood *et al.* propose learning-by-synthesis for eye shape registration and gaze estimation. They train a Constrained Local Neural Field (CLNF) [10] deformable model on synthetic eye images with annotated landmarks. The authors [11] extract discriminatory Haar features from 2D Haar wavelet transform and a new efficient Support Vector Machine (SVM) is proposed for eye detection. These learning-based methods require large amounts of training data and it can be unreliable for manual annotation especially for labeling the pupil center.

For the learning-based methods, learning-by-synthesis is a promising solution for large training data collection and has been widely used for appearance-based eye gaze estimation

[12], [9], [13]. Computer graphics techniques are applied to automatically generate large amounts of synthetic labeled photorealistic eye images. These synthetic eye images can cover various appearance respect to different gaze directions, head poses and illumination conditions. In addition, the eye related landmarks as well as pupil center can be perfectly labeled. Hence, we propose to learn from synthetic eye images for eye detection.

III. Proposed Framework

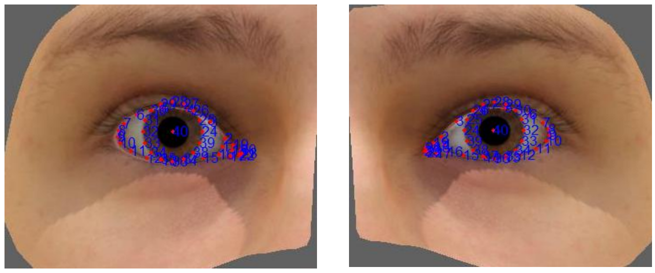
Discriminative regression-based models for facial alignment are well-studied and achieve state-of-the-art performance on facial landmark detection. In this work, motivated by facial landmark detection [14] and learning-by-synthesis for appearance-based gaze estimation [12], we propose to learn cascade regression models from synthetic data for accurate eye detection. Different from [7] which learns regression models from real labeled data and tests on given eye regions, our model learned from pure synthetic data can automatically predict the eye center on a real image. Moreover, in this work, we capture shape information of eye related landmarks for eye center detection. In this section, we firstly give a brief introduction about our synthetic training data. Then more details of our proposed approach are discussed.

A. Synthetic Eye

Existing work aim to learn appearance-based eye detector from few samples [11], [7]. Collecting large amounts of training data with ground truth of eye center and other eye landmarks locations is costly, very time-consuming and also can be unreliable. In addition, it is difficult to collect enough training samples to cover various illuminations and subjects with different head poses, eye states or gaze directions. Hence, existing learning-based eye detection methods are not general to in-the-wild scenarios since the specific collected training samples are limited.

Recently, learning from synthetic eye images for appearance-based eye gaze estimation achieves state-of-the-art results [9], [12]. They synthesize large scale variable eye images, which make it possible to cover all the appearance respect to different head poses, illumination conditions and gaze directions. In [12], the authors present a novel method named *UnityEyes* for rapidly synthesizing variable eye regions images. They derive 3D eye region model from high-resolution 3D face scans. Eyeballs are modeled using a single 3D mesh which corresponds to the eye's external surface. In addition, the shape of 3D mesh is defined by two spheres representing the cornea and the sclera. By scaling the iris boundary and texture-shape offset, variable eyeball shape and texture including iris width and dilation can be generated in *UnityEyes*. Eyelid animation is achieved through geometric methods from anatomy. Different head poses can be simulated by using spherical coordinates and pointing it towards the eyeball center of 3D model as shown in the firs row of Fig. 1. Another important factor is the illumination condition especially for the appearance based methods for

Fig. 1. Samples of synthetic eyes with different head poses (first row), various illuminations (second row) and eye appearance (third row)



(a) (b)

Fig. 2. Sample of flipped right eye (a) and corresponding original synthetic left eye (b), with 2D facial landmarks (in this work, 11 landmarks indexed by [3-6 9 12-15 18 40] are used).

eye detection. They produce highlights and soft shadows by simulated light sources pointing in a random direction towards the eye regions. They further choses panoramic photographs and randomly vary the rotation and exposure levels to render the reflections and environmental ambient light. Some synthetic images are shown in Fig. 1. Their experimental results show that even K-nearest-neighbor algorithm training on these synthetic data for eye gaze estimation outperforms other deep learning based methods training on real data. More details about *UnityEyes* can be found in [12].

In this paper, we propose to learn cascade regression models from synthetic eyes. We use the *UnityEyes* [12] to synthesize 10730 eye region images. Some synthetic training samples are shown in Fig. 1. In addition, 2D facial landmarks such as eye corners and eye center are available. Since *UnityEyes* generates eye region images of left eye, we flip them to train the right eye model (see Fig. 2). During the testing, left and right eye are located separately using corresponding model.

B. Cascade Regression for Landmark Detection from Synthesis

Although we focus on eye center detection, other contextual information such as eye corners and eyelids can be helpful for eye center detection. Cascade regression has been successfully applied for numbers of facial landmarks detection and

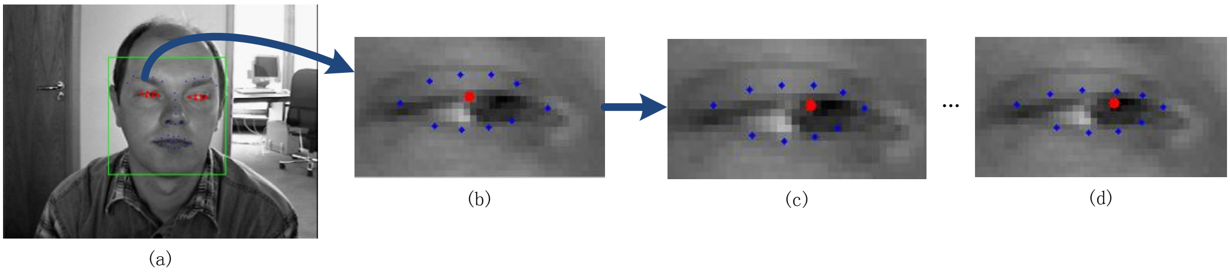


Fig. 3. Framwork of our proposed approach. (a) Face detection and 51 facial landmrk detection. (b) Crop eye region and initialize 11 points. (c) Output of first iteration. (d) Final eye landmarks detection.

achieves state-of-the-art results [8], [14], [15]. Similar to face alignment, eye related landmarks including the eye centers can be detected by cascade regression frame work. Our proposed coarse-to-fine framework for eye detection is summarized as shown in Fig. 3. We use a fast Deformable Part Models (DPM) based face detector [16] provided in [17] for face detection. Then facial landmark detection [14] is adopted to coarsely localize the eye region. After cropping out both right and left eye regions, cascade regression frame work is applied for accurate eye detection. More details are described below.

The overall framework of cascade regression for eye detection is shown in Algorithm 1. In this work, we select out 11 eye related key points (see Fig 3) consisting of 2 eye corners, 8 eyelid points and 1 pupil center. After extracting eye regions based on coarse landmark detections, the 11 key point locations $\mathbf{x} \in \Re^{2\cdot 11}$ are initialized by mean eye of the training data. Since there are various eye shapes respect to the position of eyelids and head poses, similar to [14] for facial landmark detection, we propose to iteratively update the regression parameters. In cascade regression, key points locations are iteratively updated based on the learned regression models. At each iteration t, shape features (Euclidean distance of each pair of points), local appearance features (SIFT) from the eye centers as well as contextual features (SIFT) from eyelids and eye corners are jointly used to predict the key point locations. A linear regression model f_t is used to predict the location updates $\Delta\mathbf{x}^t$ on the basis of current key point locations $\mathbf{x}^{t-1}$. f_t is defined as below:

$$f_t : \Delta\mathbf{x}^t = \alpha^t \Phi(I, \mathbf{x}^{t-1}) + \beta^t \Psi(\mathbf{x}^{t-1}) + \mathbf{c}^t \tag{1}$$

where $\Phi(I, \mathbf{x}^{t-1}) \in \Re^{11\cdot 128}$ denotes the corresponding local appearance of pupil and contextual features of other key points, and $\Psi(\mathbf{x}^{t-1}) \in \Re^{11\cdot 10}$ represents the shape features. α, β and $\mathbf{c}$ are the parameters of regression model.

For the training, given ith synthetic eye image $\mathbf{I}_i$ with 11 ground truth landmark locations $\mathbf{x}_i^*$, at tth iteration, the ground truth eye shape updates can be calculated by:

$$\Delta\mathbf{x}_i^{t,*} = \mathbf{x}_i^* - \mathbf{x}_i^{t-1} \tag{2}$$

where when t equals 1 at the first iteration, $\mathbf{x}_i^0$ is the mean locations of key points calculated from all training data. We

---

Algorithm 1 Cascade regression for eye detection.

Input:
Give the eye image $\mathbf{I}$. 11 key point locations $\mathbf{x}^0$ are initialized by mean eye.
Do cascade regression:
for t=1,...,T **do**
Estimate the key point location updates given the current key point locations $\mathbf{x}^{t-1}$.
$f_t : \mathbf{I}, \mathbf{x}^{t-1} \to \Delta\mathbf{x}^t$
Update the key point locations.
$\mathbf{x}^t = \mathbf{x}^{t-1} + \Delta\mathbf{x}^t$
end for
Output:
Acquire locations $\mathbf{x}^T$ of eye landmarks including the eye center.

---

learn the parameters of regression model by a standard least-square formulation with closed form solution:

$$\begin{aligned}\alpha^{t^*}, \beta^{t^*}, \mathbf{b}^{t^*} = & arg \min_{\alpha^t, \beta^t, \mathbf{c}^t} \sum_{i=1}^{K} \| \Delta\mathbf{x}_i^{t,*} - \alpha^t \Phi(I, \mathbf{x}^{t-1}) \\ & - \beta^t \Psi(\mathbf{x}^{t-1}) - \mathbf{c}^t \|^2\end{aligned} \tag{3}$$

Then for the testing, in this paper, we empirically set the number of iterations T as 4. At iteration t, 11 key point location updates $\Delta\mathbf{x}^t$ can iteratively be estimated according to Equation 1. Then key point locations for current iteration can be acquired though $\mathbf{x}^t = \mathbf{x}^{t-1} + \Delta\mathbf{x}^t$. $\mathbf{x}^0$ is the mean location of 11 key points from the synthetic training data.

IV. Experimental Results

A. Evaluation Database and Measurement Criteria

We evaluate the proposed eye localization method and compare it with the state-of-the-art methods on two widely used benchmark databases, including GI4E [18] and BioID [19]. GI4E [18] images have a resolution of 800×600 in pixel and are representative for the ones that can be acquired by a normal camera. It contains 1236 images of 103 subjects with 12 different gaze directions. BioID is one of the most widely used database for eye center localization which contains 1,521 gray, 23 subjects and 384×286 resolution images. This database is very challenging with different face sizes, various illuminations and subjects with glasses or closed eyes.

In the following experiments, we use the maximum normalized detection error [19] to evaluate the performance of eye detection. It is calculated as below:

$$d_{eye} = \frac{max(D_r, D_l)}{\|L_r - L_l\|} \tag{4}$$

where D_r and D_l are the Euclidean distances between the estimated right and left eye centers and the ones in the ground truth, and L_r and L_l are the true centers of the right pupil and left pupil respectively. d_{eye} is normalized by the inter-ocular distance. It measures the error obtained by the worst of both eye estimation. In this measure, $d_{eye} \le 0.25$ corresponds to the distance between eye center and eye corner, $d_{eye} \le 0.1$ corresponds to the range of iris, and $d_{eye} \le 0.05$ corresponds to the range of pupil diameter.

B. Experiments and Comparisons

The experimental results of GI4E and comparisons with other methods are listed in Table I. The best performance for evaluation criteria is highlighted in bold in Table I. Our approach achieves the best results when $d_{eye} \le 0.05(98.2\%)$ and $d_{eye} \le 0.1(99.8\%)$ on GI4E database. At $d_{eye} \le 0.25$, the accurate detection rate is close to the best one reported in [20]. Actually, we further investigate that there are two false positives for face detection which result in the detection rate of 99.8%. Although our improvement is marginal compared with the method proposed in [20], our method is more robust because it performs much better on more challenging BioID database.

The first step for face detection is very important for final eye detection. Since most of existing works test on BioID database using Viola Jones face detector [21] for face detection, to fairly compare with state-of-the-art method, we apply VJ face detector in this work. The evaluations and further comparisons with state-of-the-art eye localization methods on BioID are shown in Table II and Fig. 4. As shown in Table II, for normalized error $d_{eye} \le 0.05$, $d_{eye} \le 0.1$ and $d_{eye} \le 0.25$, our method achieves a detection rate of 89.2%, 98.0% and 99.8% respectively, which are better than state-of-the-art methods.

Figure 5 shows some examples of eye detection on GI4E and BioID database, where the white dot represents the manual annotation and red dot denotes the estimated eye location. Even though the eye is closed or subject with glasses, our proposed method can still predict the eye center locations. From the qualitative results in Fig. 5, the predicted eye center locations are even more reasonable than the manual annotation of ground truth in some cases. As shown in Fig. 6, for the first sample, it fails to detect the eye center when the eyelids and pupils are not apparent. Learned regression models fail to capture the contextual and local appearance information for accurate eye detection. Moreover, not so accurate face detection is more likely to result in failure of landmark detection, and leads to false eye center detection because of wrongly cropped eye region (see Fig. 6). It is feasible to tackle this problem by applying more accurate face detector.

TABLE I
EYE LOCALIZATION COMPARISON ON GI4E DATABASE

Method	$d_{eye} \le 0.05$	$d_{eye} \le 0.1$	$d_{eye} \le 0.25$
Timm2011 [4]	92.4%	96.0%*	97.5%*
Villanueva2013 [18]	93.9%	97.3%*	98.5%*
Zhang2016 [20]	97.9%	99.6%	**99.9%**
Ours	**98.2%**	**99.8%**	99.8%

* are estimated from the accuracy curves in corresponding paper [18].

TABLE II
EYE LOCALIZATION COMPARISON ON BIOID DATABASE

Method	$d_{eye} \le 0.05$	$d_{eye} \le 0.10$	$d_{eye} \le 0.25$
Campadelli2009 [22]	80.7%	93.2%	99.3%
Timm2011 [4]	82.5%	93.4%	98.0%
Valenti2012 [3]	86.1%	91.7%	97.9%
Araujo2014 [5]	88.3%	92.7%	98.9%
Ren2014 [23]	77.1%	92.3%	99.0%
Chen2015 [11]	88.8%	95.2%	99.0%
Zhang2016 [20]	85.7%	93.7%	99.2%
Ours	**89.2%**	**98.0%**	**99.8%**

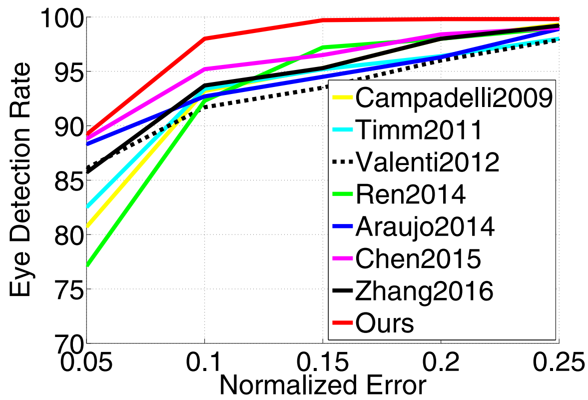


Fig. 4. Eye detection rate of the proposed method on the BioID database, in comparison with other state-of-the-art methods.

C. Further Discussion

We also apply the face detector [17] for testing on BioID, which is the same detector as tests on GI4E. It achieves a final eye detection rate of 87.6% of $d_{eye} \le 0.05$ on BioID. It is lower than the results reported in Table II because for low resolution images, face detection affects more on eye region extraction after normalizing the eye width to 25 in pixel. By further investigation, the facial landmark detection model is trained from VJ face detector which affects the facial landmark detection. In a word, face detection affects initialization of eye landmarks which is important for final eye detection especially in low resolution images. In this paper, false negatives of

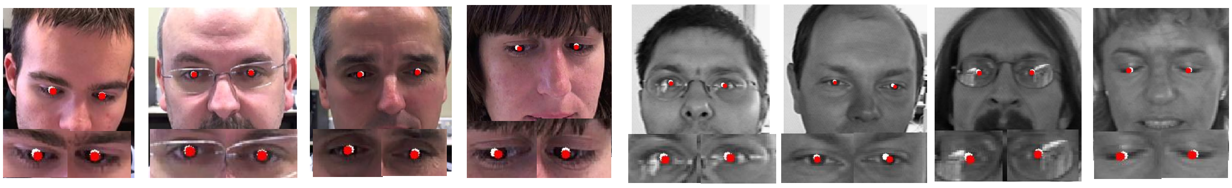


Fig. 5. Eye deteciton examples of successes on GI4E and BioID database. The white dot represents the grund truth and red dot represents estimated eye locations.

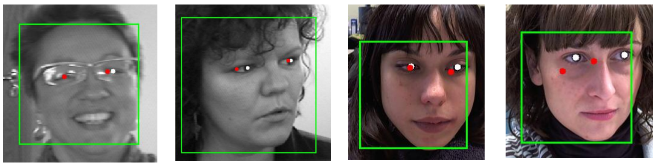


Fig. 6. Eye detection examples of failure on BioID (first two samples) and GI4E (last two samples) database. The white dot represents the grund truth, red dot represents estimated eye locations and green bounding box is detected face (best view in color).

face detection are discarded similar to other literatures. Face detection rates are 97.5% and 99.8% on BioID and GI4E database respectively.

As discussed aforementioned, the proposed framework for eye detection is sensitive to eye region detection since the initialization of key points is very important. If the initialization is far away from the ground truth, it can not converge to the global optimization. Another existing method [7] based on SDM [8] do the eye center detection on given eye region images of BioID. We do a similar experiments and crop the eye region based on ground truth landmark locations instead of automatically detection of eye related landmarks. Here we achieve a detection rate of 94.2% on BioID database which is better than 93.8% reported in [7] at normalized error $d_{eye} \leq 0.05$. It should be noted that our model learn from pure synthetic data while they train from more than 10,000 real images with manual annotations.

In our experiments, we first generate 2218 training images by *UnityEyes* and acquire a detection rate of 94.2% on GI4E dataset. It is not preferable for real application. By further investigation, we find out that the generated synthetic images is limited in covering the variance for our learning based methods. To cover the variance of most realistic eyes with different poses and illuminations, we learn our model from 10730 synthetic eye images which are randomly generated and get a detection rate of 98.2%. However, pure synthetic data has some disadvantages. *UnityEyes* is not possible to model subjects with glasses which make it clearly lacks some variance that may occur in natural real data. We further train a model form combination of synthetic data and real data (BioID) using 5 eye landmarks and test on GI4E. It only achieve a detection rate of 95.4% which performs worse than training from synthetic data with a detection rate of 98.2% at normalized error $d_{eye} \leq 0.05$. The reasons come that only 5 eye landmarks are available which captures less contextual information and annotations of real data are not all very accurate. Further works will focus on theses experiments.

Our contributions over a similar work presented in [9] are in three folds. First, compared with the work presented in [9] which focus on learning-by-synthesis for gaze estimation, we contribute to eye detection and conduct quantitative and qualitative experiments on benchmark datasets. Second, we utilize the cascade regression framework on 11 landmarks to boost the performance and efficiency while they achieve eye shape registration by training separate CLNF patch experts to fit the 28 landmarks. Moreover, except for capturing the local appearance features, we incorporate the shape features for eye center locations. Their public available method is provided without trained model for iris registration. We will implement the training method and further conduct experiments with [9] for comparisons.

All experiments are conducted with nonoptimized Matlab codes on a standard PC, which has an Intel i5 3.47GHz CPU. 15 frames per second on BioID can be achieved when applying Viola Jones face detector in our proposed framework, which allows for near real time eye detection.

V. Conclusions

In this paper, we present a novel and effective framework for accurate eye detection based on the cascade regression model learned from pure synthetic data. Coarse-to-fine strategy is employed. Facial landmarks are firstly detected and then the eye related detected landmarks are used to localize the eye regions. Cascade regression starts from the initial eye shape at eye region and iteratively updates the eye related key points based on the regression models learned from synthetic eye images. Experimental results show that our method outperforms state-of-the-art methods for eye detection.

Future work will focus on its extensive applications like gaze estimation and eye tracking. In addition, learning from appropriate combination of synthetic data and real data will potentially improve the results.

Acknowledgments

This work was completed when the first author visited Rensselaer Polytechnic Institute (RPI), supported by a scholarship from University of Chinese Academy of Sciences (UCAS). The authors would like to acknowledge support from UCAS and RPI. This work was also supported in part by National Science Foundation under the grant 1145152 and by the National Natural Science Foundation of China under Grant 61304200 and 61533019.

REFERENCES

[1] D. W. Hansen and Q. Ji, "In the eye of the beholder: A survey of models for eyes and gaze," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 3, pp. 478–500, 2010.

[2] F. Song, X. Tan, S. Chen, and Z.-H. Zhou, "A literature survey on robust and efficient eye localization in real-life scenarios," *Pattern Recognition*, vol. 46, no. 12, pp. 3157–3173, 2013.

[3] R. Valenti and T. Gevers, "Accurate eye center location through invariant isocentric patterns," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 9, pp. 1785–1798, 2012.

[4] F. Timm and E. Barth, "Accurate eye centre localisation by means of gradients." in *VISAPP*, 2011, pp. 125–130.

[5] G. Araujo, F. Ribeiro, E. Silva, and S. Goldenstein, "Fast eye localization without a face model using inner product detectors," in *Image Processing (ICIP), 2014 IEEE International Conference on*, Oct 2014, pp. 1366–1370.

[6] N. S. Karakoc, S. Karahan, and Y. S. Akgul, "Regressor based estimation of the eye pupil center," in *Pattern Recognition*. Springer, 2015, pp. 481–491.

[7] M. Zhou, X. Wang, H. Wang, J. Heo, and D. Nam, "Precise eye localization with improved sdm," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 4466–4470.

[8] X. Xiong and F. De la Torre, "Supervised descent method and its applications to face alignment," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 532–539.

[9] E. Wood, T. Baltrušaitis, X. Zhang, Y. Sugano, P. Robinson, and A. Bulling, "Rendering of eyes for eye-shape registration and gaze estimation," *arXiv preprint arXiv:1505.05916*, 2015.

[10] T. Baltrusaitis, P. Robinson, and L.-P. Morency, "Constrained local neural fields for robust facial landmark detection in the wild," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2013, pp. 354–361.

[11] S. Chen and C. Liu, "Eye detection using discriminatory haar features and a new efficient svm," *Image and Vision Computing*, vol. 33, pp. 68–77, 2015.

[12] E. Wood, T. Baltrušaitis, L.-P. Morency, P. Robinson, and A. Bulling, "Learning an appearance-based gaze estimator from one million synthesised images," in *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications*. ACM, 2016, pp. 131–138.

[13] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling, "Appearance-based gaze estimation in the wild," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4511–4520.

[14] Y. Wu and Q. Ji, "Shape augmented regression method for face alignment," in *Proceedings of the IEEE International Conference on Computer Vision Workshops*, 2015, pp. 26–32.

[15] ——, "Constrained joint cascade regression framework for simultaneous facial action unit recognition and facial landmark detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[16] M. Mathias, R. Benenson, M. Pedersoli, and L. Van Gool, "Face detection without bells and whistles," in *Computer Vision–ECCV 2014*. Springer, 2014, pp. 720–735.

[17] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," *Proceedings of the British Machine Vision*, vol. 1, no. 3, p. 6, 2015.

[18] A. Villanueva, V. Ponz, L. Sesma-Sanchez, M. Ariz, S. Porta, and R. Cabeza, "Hybrid method based on topography for robust detection of iris center and eye corners," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 9, no. 4, p. 25, 2013.

[19] O. Jesorsky, K. J. Kirchberg, and R. W. Frischholz, "Robust face detection using the hausdorff distance," in *Audio-and video-based biometric person authentication*. Springer, 2001, pp. 90–95.

[20] W. Zhang, M. L. Smith, L. N. Smith, and A. Farooq, "Eye center localization and gaze gesture recognition for human–computer interaction," *JOSA A*, vol. 33, no. 3, pp. 314–325, 2016.

[21] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.

[22] P. Campadelli, R. Lanzarotti, and G. Lipori, "Precise eye and mouth localization," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 23, no. 03, pp. 359–377, 2009.

[23] Y. Ren, S. Wang, B. Hou, and J. Ma, "A novel eye localization method with rotation invariance," *Image Processing, IEEE Transactions on*, vol. 23, no. 1, pp. 226–239, 2014.