2018 IEEE Games, Entertainment, Media Conference (GEM).

A Deep Learning approach to Segmentation of Distorted Iris regions in Head-Mounted Displays

Viktor Varkarakis
Collage of Engineering and Informatics National University of Ireland Galway
Galway, Ireland
Email: v.varkarakis1@nuigalway.ie

Shabab Bazrafkan
Collage of Engineering and Informatics National University of Ireland Galway
Galway, Ireland
Email: s.bazrafkan1@nuigalway.ie

Peter Corcoran
Collage of Engineering and Informatics National University of Ireland Galway
Galway, Ireland
Email: peter.corcoran@nuigalway.ie

***Abstract*—In this paper, we consider the next generation of wearable AR/VR display glasses and the challenges of personal authentication on such devices. The use of iris authentication as a mean of creating a seamless biometric link between the user and his personal data offers a viable approach, but due to the likely location of user-facing cameras there are some challenges in achieving an accurate segmentation of the iris. In this paper, a deep neural network was trained to accurately segment distorted iris regions. An appropriate augmentation method is presented to generate the distorted iris dataset used for training from publicly available frontal iris datasets. The proposed method shows promising results in segmenting off-axis iris images in unconstrained conditions.**

I. Introduction

Virtual reality display technology has only recently appeared as a consumer electronics product in the form of new VR headsets, but in industry sectors these headsets have been evolving for more than a decade [1]. But now that the technology is ready for mass market deployment it will not take long for the next generation of headset technology to evolve, providing even more sophisticated display and interface capabilities and acting as a gateway into new virtual and augmented application frameworks [2]–[5].

Now the current evolution in AR display systems does not imply that this is a ‘new’ technology. Researchers have in fact been working with Augmented Displays for more than 20 years [6]–[12]. The Microsoft HoloLens [3], [13]–[15] is a good example of an AR headset that is similar to today’s VR headsets – perhaps a little oversized for day-to-day consumer use. The most recent mass market experiment with a wearable, augmented/mediated-reality display, that could be worn on a day-to-day basis, was Google Glass[11],[16], [17]. Glass, as it became known, was considered to be a game changing technology for a few years across a wide range of industry sectors [18]–[21]. Ultimately, however, the product was withdrawn [22] and since that point consumers have had to wait for the next mass-market AR display technology.

Currently there are several companies that are working on the next generation of wearable technology to follow in the footsteps of Google Glass and Steve Mann’s EyeTap. There are numerous challenges both in the display technology itself, but also in the ergonomics and the user-interface aspects of the device. Nevertheless, we expect that recent advances in motion-sensing technologies, eye-tracking and affective interpretation models will improve the usability of the next generation of these devices.

Another key challenge which applies to all variants of AR and VR headsets is that of user authentication. Where there is no physical keyboard or equivalent there are challenges to authenticate a user and ensure the privacy of user profiles and data that is collected/generated by a wearable AR consumer device. It is interesting that in the recently released movie “Ready Player One” the players who enter the virtual world of “The Oasis” are portrayed as being linked auto-magically with their virtual world profiles, but no explanation of this magic “authentication” achieved, is provided. Fortunately, we can answer this question for you in this short paper by looking at how AR glasses are evolving today and demonstrating how current biometric authentication technology can be easily re-purposed to work in a seamless manner on tomorrow’s wearable AR/VR glasses and display devices.

A good starting point is found in [23] where the author considers how biometric technology is becoming the natural authentication mechanism for personal consumer devices and the broad range of services and capabilities they bring to our daily lives. In follow-on publications the use of iris recognition on consumer devices is explored [24]–[27] and the importance of accurate iris segmentation, particularly in consumer imaging devices, is identified as a key challenge [28], [29]. In the iris authentication workflow, failed segmentations represent the single largest source of error [30]–[32].

In addition to its role in improving the performance of an iris-based authentication system, the accurate segmentation of iris regions, can be used successfully for eye-gaze estimation [33]. Eye-gaze is a key element of various user-interface modalities for wearable AR & VR displays.

As shown in US design patent, D795952S1 [34] “Fig1”, a possible location of the camera used for iris authentication or eye-gaze direction, is below the eye. As opposed to when the camera is frontal, resulting in having circular iris images, in positioning the camera, same way or similar as described above, off-axis iris images will be obtained. At the moment

Authorized licensed use limited to: Yunnan University. Downloaded on November 07,2025 at 15:57:24 UTC from IEEE Xplore. Restrictions apply.

there is not a publicly available commercial segmentation method for off-axis iris images as this is quite a new problem arising from the recent proliferation of AR/VR technology into consumer devices.

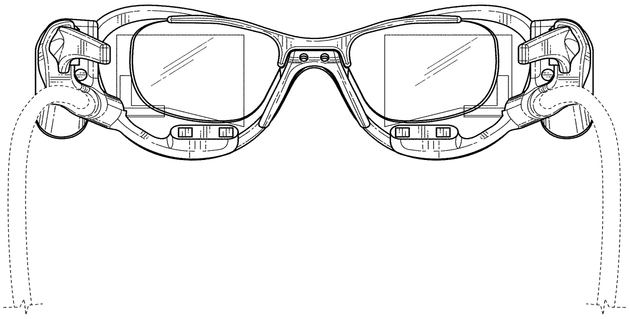

Fig. 1. Virtual Reality Glasses - Patent Number: USD795952S1 [34]

The focus of this work is on implementing a deep learning technique to segment distorted iris images. The main contribution is a data augmentation technique that simulates iris images from an AR/VR head mounted display with off-axis cameras. On a selection of the distorted iris images a second augmentation process is applied, adding contrast, blurring or shadows to the images mimicking the variability of images quality achieved in real-life environment.

In the next section the proposed method is explained in more detail including the data preparation and augmentation techniques that are implemented. The design of the deep learning network is also presented along with a description of the underlying training algorithm. The results of the proposed method are given in the final section.

II. Proposed Method

A. Database Preparation

In order to accurately train a deep neural network, a large number of labeled training samples are required. These samples should correctly characterize the imaging problem so that it can be solved, enabling the deep learning process to train an accurate model. For this reason, an augmentation process was implemented and is explained in section: B. Data Augmentation. The CASIA Thousand [35] and Bath 800 [36] databases were used as a starting point for the task of generating a suitably augmented training dataset. The CASIA Thousand has approximately 20.000 images and Bath 800 over 30.000 iris images. These databases are not labeled with the segmentation ground truth, but they do contain high quality iris images (captured in constrained conditions). There are high resolution, high contrast samples, with low noise and shadowing. These high-quality images can be accurately segmented with standard industrial segmentation algorithms applied to segment the original images. In this work, these segmentations are considered as the effective ground truth for the iris images.

B. Data Augmentation

To proceed to the training of a deep neural network for AR/VR iris segmentation task, an augmentation of the database is required as there is not a publicly available database of iris images acquired from an AR/VR set-up. Therefore, the first objective of the augmentation process is to simulate the representation of respective iris images. This representation, consist of distorted iris images, some with secondary low contrast, blurring and shadowing of the distorted iris images. In this work, all the samples are resized to 120x160 using bicubic interpolation. To generate the initial distorted iris images the samples were warped using two transformations.

1. The images are warped by applying a spatial stretching/contacting of the iris images in different parts. For example, an image will be stretched in its left side and contracted in its right side as shown in “Fig. 2”. The warping is applied in a linear manner with random parameters. The parameters determine the amount of stretching/contracting of the image in each direction. The void spaces are filled using linear interpolation. This transformation is applied in both horizontal and vertical direction. After that the images are resized to the original resolution (120x160). Applying warping results in iris images with non-circular iris and pupil structures, as shown in “Fig.3” which is a usual case in iris images acquired from an AR/VR headset. This transformation is applied with 50% probability to the images.

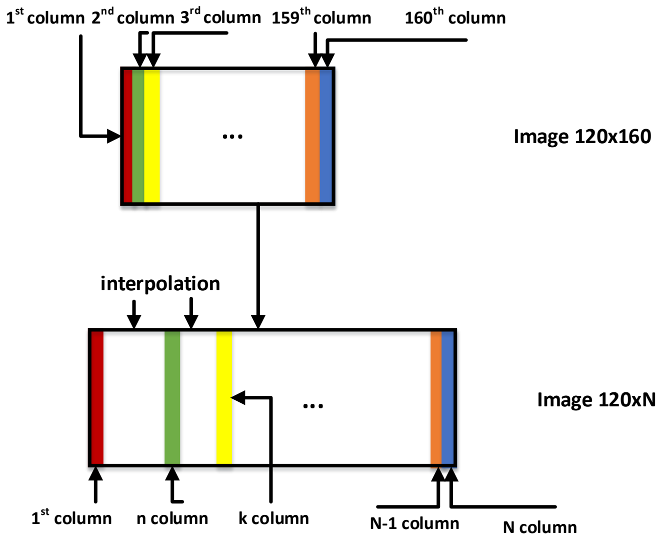

Fig. 2. Stretching/Contracting of the image.

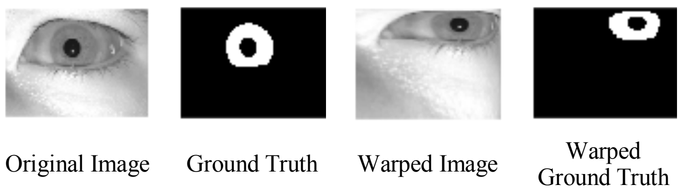

Fig. 3. Example of a warped image.

2. At the second transformation, the images are tilted in two directions (up-left, up-right) with the same probability.

This research is funded under the SFI Strategic Partnership Program by Science Foundation Ireland (SFI) and FotoNation Ltd. Project ID: 13/SPP/I2868 on Next Generation Imaging for Smartphone and Embedded Platforms.

Portions of the research in this paper use the CASIA-IrisV4 collected by the Chinese Academy of Sciences' Institute of Automation (CASIA).

Authorized licensed use limited to: Yunnan University. Downloaded on November 07,2025 at 15:57:24 UTC from IEEE Xplore. Restrictions apply.

This is accomplished by applying a projective transformation to the images. This transformation is mapping the top vertices of the image to a new pair of point as illustrated in "Fig. 4". In "Fig. 4" the values of a, b, c and d are randomly generated so that the image is tilted in the desired direction. The values from "Fig.4", of a and c are in U (0.9,1), b is in U (0.15,0.45) and d in U (0.55,0.85) where U is the uniform distribution. The images were tilted in these directions, due to the fact that, as shown in "Fig. 1" and described in the introduction, a possible location of the camera used for capturing the iris images will be below the eye. The second transformation is applied to the images that the first transformation was not applied, along with the images that were distorted in the previous step with 50% probability. In "Fig.5" an image is shown where the above transformation was applied.

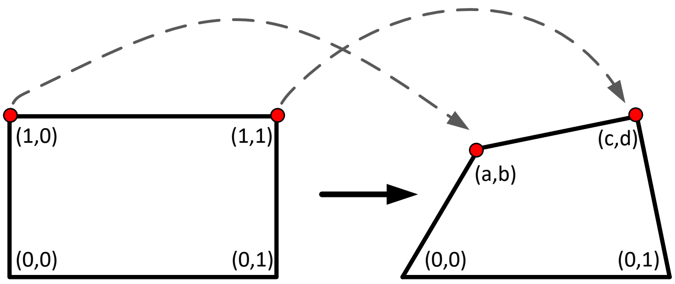

Fig. 4. How the projective transformation is applied to an image.

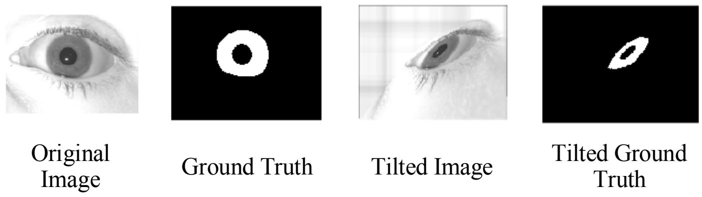

Fig. 5. Example of tilted image.

The second objective of the augmentation, is to ensure that the samples used to train the network represent real-life scenarios. The distribution of the input data plays an extremely important role in what the network learns and how will behave during the testing stage but also in unconstrained situations [37]. The databases that are used, CASIA Thousand and Bath 800, consist of high quality iris images. Based on the precise observations that have been done [26], [38] the main differences between high quality constrained images and wild ones, are linked to contrast, blurring and shadows in the image. In order to simulate real-life captured iris images, the contrast inside and outside the iris region is changed separately using histogram mapping. Motion blurring is applied, as well as, adding shadows to the images by multiplying them with a shadow function. Contrast and blurring are applied to all the images and with 50% probability shadowing is applied to the samples. The functions used in order to apply contrast, shadowing and motion blurring are explained in [29]. The augmentation process is shown in "Fig. 6". The distortion-contrast database and the distortion database, as illustrated in "Fig. 6", are created twice and the contrast database once.

After the augmentation process, with the addition of the original images the total number of samples generated is over 300.000. Regarding the ground truth of the augmented data, when a distortion method is applied to an image, an identical distortion is also applied to its ground truth segmentation. In "Fig.7" examples of augmented data are presented along with their corresponding augmented ground truth.

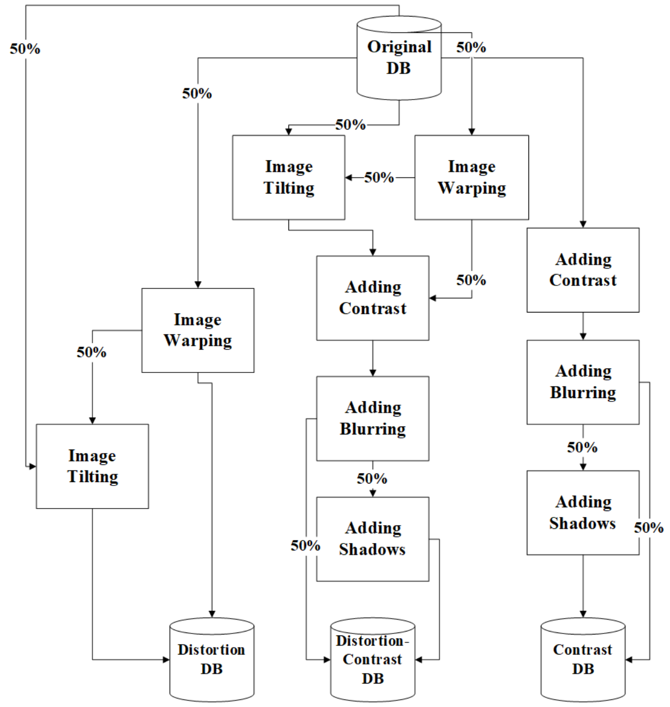

Fig. 6. Workflow of the augmentation process.

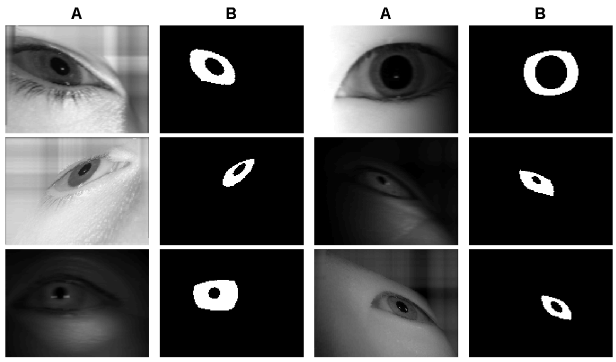

Fig. 7. A: Augmented samples, B: Augmented Ground Truth of Segmentation

C. Network Design

For the segmentation task a fully convolutional network inspired by [29] is used, consisting of 10 layers. The network starts with a 3x3 kernel mapping the input (1 channel) on the first convolutional hidden layer which consists of 32 channels using a rectified linear unit (ReLu) as an activation function. The kernel size remains the same throughout the convolutional hidden layers, as well as, the number of channels and their activation function. Finally, at the output layer (1 channel), the kernel size is 3x3, but in this layer, the sigmoid activation function is used. Pooling layers were not used as it was

Authorized licensed use limited to: Yunnan University. Downloaded on November 07,2025 at 15:57:24 UTC from IEEE Xplore. Restrictions apply.

observed that the accuracy of the network's output was decreasing.

D. Training

The training was carried out in TensorFlow library. The Mean Squared Error is used as the loss function. The Gradient Descent with Adaptive Moment Estimation (Adam) is used, with a learning rate of 1e-4, beta1 and beta2 equal to 0.9 and 0.999 respectively, in order to optimize the loss function. The samples are divided 70% for the training set, 20% for validation set and 10% for test set. The training is done on a desktop computer with Nvidia GTX 1080 GPU.

III. RESULTS AND DISCUSSION

In this work, the network is trained on the original images and the augmented databases (Bath 800 and CASIA Thousand). The output of the network is a grayscale segmentation map with values between 0 and 1. The binary map is obtained by using a thresholding technique, where the values bigger than the threshold are shifted to 1 and the others to 0. The threshold value 0.55 is used in this work. In "Fig.8 the output of the network is presented for several sample images. Several metrics have been used to evaluate the network. The metrics used are described thoroughly in [29].

The segmentation results for the test set on the two databases (Bath 800 and CASIA Thousand) are presented below in Table I and II. In Table I, higher performance is represented by higher values and in Table II higher performance is represented by lower values.

The proposed method shows promising results using the deep learning technique in segmenting off-axis iris images as represented by AR/VR set-ups, including also effects on the images from real-life environments.

This work is an initial proof-of-concept and more details regarding the augmentation process along with numerical analysis and comparisons with other segmentation algorithms will be presented at the conference.

TABLE I. SEGMENTATION RESULTS

Metrics	***Proposed Method***	
	BATH 800	***CASIA Thousand***
Accuracy	99.12%	99.34%
Sensitivity	92.75%	90.32%
Specificity	99.58%	99.78%
Precision	94.26%	94.78%
NPV	99.45%	99.52%
F1-Score	93.23%	92.07%
MCC	92.94%	92.02%
Informedness	92.34%	90.10%
Markedness	93.72%	94.30%

TABLE II. SEGMENTATION RESULTS

Metrics	***Proposed Method***	
	BATH 800	***CASIA Thousand***
FPR	0.41%	0.22%
FNR	7.24%	9.67%
FDR	5.73%	5.21%

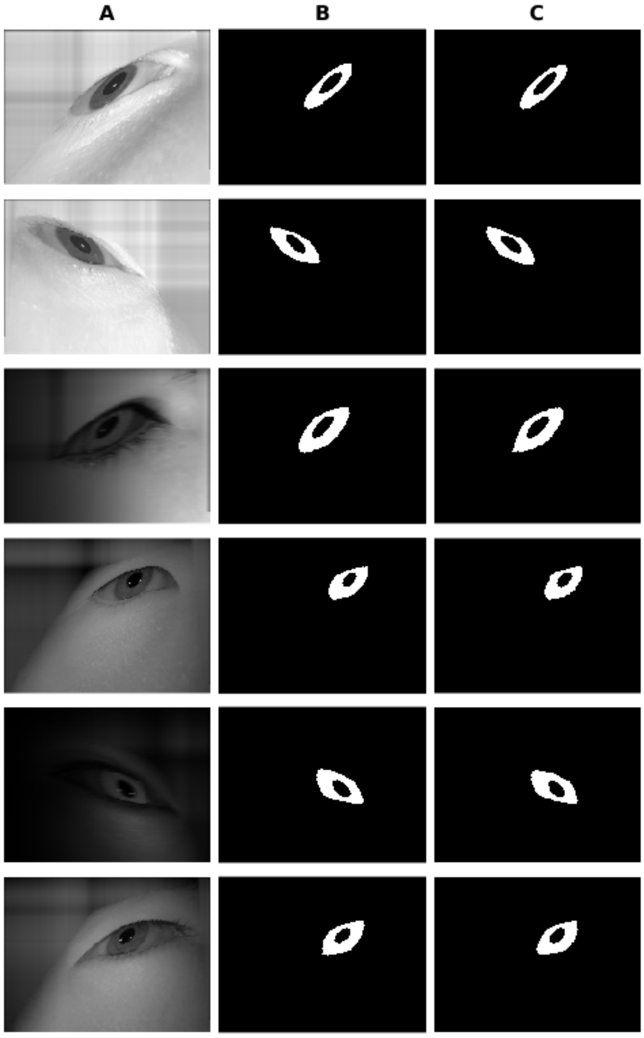

Fig. 8. A: Augmented samples, B: Segmentation Ground Truth, C: Output of Network - Segmentation map

REFERENCES

[1] B. Kress, E. Saeedi, and V. Brac-de-la-Perriere, "The segmentation of the HMD market: optics for smart glasses, smart eyewear, AR and VR headsets," in *Photonics Applications for Aviation, Aerospace, Commercial, and Harsh Environments V*, 2014, vol. 9202, p. 92020D.

[2] D. De Angeli and E. J. O'Neill, "Development of an Inexpensive Augmented Reality (AR) Headset," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '15*, 2015, pp. 971–976.

Authorized licensed use limited to: Yunnan University. Downloaded on November 07,2025 at 15:57:24 UTC from IEEE Xplore. Restrictions apply.

[3] R. Furlan, "The future of augmented reality: Hololens - Microsoft's AR headset shines despite rough edges [Resources-Tools and Toys]," *IEEE Spectr.*, vol. 53, no. 6, p. 21, 2016.

[4] M. Linao, "The Present And Future Of VR/AR: Applications In Education, Gaming, Commerce, And Industry," *CB Insights*, 2016.

[5] D. Wagner and D. Shmalstieg, "Making AR practical on Mobile Phones, Part 1," *IEEE Computer Graphics and Applications*, pp. 12–15, 2009.

[6] T. Starner *et al.*, "Augmented Reality through Wearable Computing," *Presence Teleoperators Virtual Environ.*, vol. 6, no. 4, pp. 386–398, 1997.

[7] S. Mann, "Fundamental issues in mediated reality, WearComp, and camera-based augmented reality," *Fundam. Wearable Comput. Augment. Reality, Lawrence Erlbaum Assoc. Inc*, pp. 295–328, 2001.

[8] S. Mann and J. Fung, "EyeTap Devices for Augmented, Deliberately Diminished, or Otherwise Altered Visual Perception of Rigid Planar Patches of Real-World Scenes.," *Presence Teleoperators Virtual Environ.*, vol. 11, no. 2, pp. 158–175, 2002.

[9] S. Mann, "Continuous lifelong capture of personal experience with EyeTap," *Proc. 1st ACM Work. Contin. Arch. Retr. Pers. Exp. - CARPE'04*, pp. 1–21, 2004.

[10] F. Tang, C. Aimone, J. Fung, A. Marjan, and S. Mann, "Seeing eye to eye: A shared mediated reality using EyeTap devices and the VideoOrbits gyroscopic head tracker," in *Proceedings - International Symposium on Mixed and Augmented Reality, ISMAR 2002*, 2002, pp. 267–268.

[11] S. Mann, "Steve Mann: My 'Augmented' Life," *IEEE Spectr.*, pp. 1–6, 2013.

[12] G. Bhorkar, "A Survey of Augmented Reality Navigation," *Found. Trends® Human–Computer Interact.*, vol. 8, no. 2–3, pp. 73–272, 2017.

[13] M. Fitzsimmons, "Hands on: Microsoft HoloLens review," *techradar*, 2016. [Online]. Available: http://www.techradar.com/reviews/wearables/microsoft-hololens-1281834/review.

[14] M. Garon, P.-O. Boulet, J.-P. Doironz, L. Beaulieu, and J.-F. Lalonde, "Real-Time High Resolution 3D Data on the HoloLens," in *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, 2016, pp. 189–191.

[15] H. Chen, A. S. Lee, M. Swift, and J. C. Tang, "3D Collaboration Method over HoloLens™ and Skype™ End Points," in *Proceedings of the 3rd International Workshop on Immersive Media Experiences - ImmersiveME '15*, 2015, pp. 27–30.

[16] E. Ackerman, "Google gets in your face: Google Glass offers a slightly augmented version of reality," *IEEE Spectr.*, vol. 50, no. 1, pp. 26–29, 2013.

[17] A. Hayes, "My Journey into Glass: Talking about Google Glass with stakeholders in the Glass Explorer Program," *IEEE Consum. Electron. Mag.*, vol. 5, no. 1, pp. 102–105, 2016.

[18] B. Fox and B. Felkey, "Potential Uses of Google Glass in the Pharmacy," *Hosp. Pharm.*, vol. 48, no. 9, pp. 783–784, 2013.

[19] O. J. Muensterer, M. Lacher, C. Zoeller, M. Bronstein, and J. Kübler, "Google Glass in pediatric surgery: An exploratory study," *Int. J. Surg.*, vol. 12, no. 4, pp. 281–289, 2014.

[20] B. Elise, "Google Glass a Game-changing Application in the Realm of Cultural Tourism," *Business Wire*, 2014.

[21] M. H. Schreinemacher, M. Graafland, and M. P. Schijven, "Google glass in surgery," *Surgical Innovation*, vol. 21, no. 6. pp. 651–652, 2014.

[22] A. Cave, "Why Google Glass Flopped," *Forbes*, 2015. [Online]. Available: http://www.forbes.com/sites/andrewcave/2015/01/20/a-failure-of-leadership-or-design-why-google-glass-flopped/#24c650d3556a.

[23] P. M. Corcoran, "Biometrics and Consumer Electronics: A Brave New World or the Road to Dystopia?," *IEEE Consum. Electron. Mag.*, vol. 2, no. 2, pp. 22–33, 2013.

[24] P. Corcoran, P. Bigioi, and S. Thavalengal, "Feasibility and design considerations for an iris acquisition system for smartphones," in *IEEE International Conference on Consumer Electronics - Berlin, ICCE-Berlin*, 2015, vol. 2015–Febru, no. February, pp. 164–167.

[25] S. Thavalengal, I. Andorko, A. Drimbarean, P. Bigioi, and P. Corcoran, "Proof-of-concept and evaluation of a dual function visible/NIR camera for iris authentication in smartphones," *IEEE Trans. Consum. Electron.*, vol. 61, no. 2, pp. 137–143, May 2015.

[26] S. Thavalengal, P. Bigioi, and P. Corcoran, "Iris authentication in handheld devices-considerations for constraint-free acquisition," *IEEE Trans. Consum. Electron.*, vol. 61, no. 2, pp. 245–253, May 2015.

[27] S. Thavalengal and P. Corcoran, "User Authentication on Smartphones: Focusing on iris biometrics.," *IEEE Consum. Electron. Mag.*, vol. 5, no. 2, pp. 87–93, 2016.

[28] S. Thavalengal, P. Bigioi, and P. Corcoran, "Efficient segmentation for multi-frame iris acquisition on smartphones," in *2016 IEEE International Conference on Consumer Electronics, ICCE 2016*, 2016, pp. 198–199.

[29] S. Bazrafkan, S. Thavalengal, and P. Corcoran, "An End to End Deep Neural Network for Iris Segmentation in Unconstraint Scenarios," *arXiv Prepr. arXiv1712.02877*, 2017.

[30] H. Hofbauer, F. Alonso-Fernandez, J. Bigun, and A. Uhl, "Experimental analysis regarding the influence of iris segmentation on the recognition rate," *IET Biometrics*, vol. 5, no. 3, pp. 200–211, 2016.

[31] H. Proença and L. A. Alexandre, "Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage," *Image Vis. Comput.*, vol. 28, no. 1, pp. 202–206, 2010.

[32] M. Erbilek, M. C. Da Costa-Abreu, and M. Fairhurst, "Optimal configuration strategies for iris recognition processing," 2012.

[33] Z. Hammal, C. Massot, G. Bedoya, and A. Caplier, "Eyes Segmentation Applied to Gaze Direction and Vigilance Estimation BT - Pattern Recognition and Image Analysis," 2005, pp. 236–246.

[34] S. Natsume, "Virtual reality glasses." Google Patents, 29-Aug-2017.

[35] "CASIA Iris Image Database." [Online]. Available: http://biometrics.idealtest.org/.

[36] S. Rakshit, "Novel methods for accurate human iris recognition," *Univ. Bath*, 2007.

[37] S. Bazrafkan, T. Nedelcu, P. Filipczuk, and P. Corcoran, "Deep learning for facial expression recognition: A step closer to a smartphone that knows your moods," in *2017 IEEE International Conference on Consumer Electronics (ICCE)*, 2017, pp. 217–220.

[38] S. Thavalengal, P. Bigioi, and P. Corcoran, "Evaluation of combined visible/NIR camera for iris authentication on smartphones," in *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2015, pp. 42–49.

Authorized licensed use limited to: Yunnan University. Downloaded on November 07,2025 at 15:57:24 UTC from IEEE Xplore. Restrictions apply.