



Deep neural network and data augmentation methodology for off-axis iris segmentation in wearable headsets

Viktor Varkarakis ^{a,*}, Shabab Bazrafkan ^b, Peter Corcoran ^a

^a Department of Electronic Engineering, College of Engineering, National University of Ireland Galway, University Road, Galway, Ireland

^b imec-Vision Lab, Department of Physics, University of Antwerp, Antwerp, Belgium



ARTICLE INFO

Article history:

Received 27 February 2019

Received in revised form 2 July 2019

Accepted 25 July 2019

Available online 1 August 2019

Keywords:

Deep neural networks

Data augmentation

Off-axis

Iris segmentation

AR/VR

ABSTRACT

A data augmentation methodology is presented and applied to generate a large dataset of off-axis iris regions and train a low-complexity deep neural network. Although of low complexity the resulting network achieves a high level of accuracy in iris region segmentation for challenging off-axis eye-patches. Interestingly, this network is also shown to achieve high levels of performance for regular, frontal, segmentation of iris regions, comparing favourably with state-of-the-art techniques of significantly higher complexity. Due to its lower complexity this network is well suited for deployment in embedded applications such as augmented and mixed reality headsets.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Data augmentation is a common technique in Deep Learning and is frequently exploited by researchers in order to overcome the obstacle of limited labelled data. But even where large datasets are available using appropriate data augmentation techniques can improve the distribution of training samples and reduce overfitting during training. In turn this increases the generalization of the trained network and improves network accuracy and robustness.

Commonly used data augmentation techniques involve rotation, translation, flipping, re-sizing or affine transformation of data samples. Other well known techniques include adding noise, motion or optical blur, or varying image contrast or gamma. These techniques are computationally inexpensive (Krizhevsky, Sutskever, & Hinton, 2012) and have been previously used successfully to reduce overfitting in training a CNN for the *ImageNet Large-Scale Visual Recognition Challenge* (ILSVRC) (Russakovsky et al., 2015), and achieved state-of the-art results at that time.

Data augmentation methods have been widely used in deep learning, and the selection of appropriate data augmentation strategies can be more important to solving a machine learning problem than the choice of a particular neural-network structure (Goodfellow, Bengio, & Courville, 2016; Shijie, Ping, Peiyi, & Siping, 2017).

In this work we focus on presenting and validating a new data-augmentation technique targeted to optimize the performance of off-axis iris segmentation. Our results show promising levels of accuracy for off-axis segmentation and the resulting trained neural network has performance that is competitive with state-of-art frontal iris segmentation networks of much greater complexity. This illustrates the powerful generalizing capabilities of our augmentation methodology.

The core of this work takes high-resolution frontal iris datasets as a starting point for generating a corresponding set of off-axis data samples. A similar approach can be adopted and applied to create a set of off-axis data samples from any frontal object view, but in this work we focus on the problem of off-axis iris segmentation as this is a new, emerging challenge for user authentication on Augmented and Mixed reality headsets.

Biometric user authentication is available on consumer devices, including smartphones, using facial recognition (Darwais, Moradian, Rahmani, & Knauer, 2014; Samangouei, Patel, & Chellappa, 2017; Vazquez-Fernandez & Gonzalez-Jimenez, 2016) and fingerprint biometric (Bakir, Chesler, & Torriente, 2016; Cherapau, Muslukhov, Arachchilage, & Beznosov, 2015; De Luca, Hang, von Zezschwitz, & Hussmann, 2015; Goode, 2014; Ring, 2015; Tipton, White II, Sershon, & Choi, 2014). The broad adoption of biometrics on consumer devices was originally discussed in Corcoran (2013) with additional discussion of the impacts in several following articles (Corcoran, 2016, 2017; Corcoran & Costache, 2016). Being a near ideal biometric, the iris of the human eye is well-suited to many consumer applications, but iris recognition is traditionally implemented in a controlled environment and under constrained acquisition conditions.

* Corresponding author.

E-mail addresses: v.varkarakis1@nuigalway.ie (V. Varkarakis), shabab.bazrafkan@uantwerpen.be (S. Bazrafkan), peter.corcoran@nuigalway.ie (P. Corcoran).

Authentication requirements in consumer devices are evolving beyond today's mobile devices. New virtual reality (VR) and augmented reality (AR) headsets provide a gateway to sophisticated virtual worlds and online services (Kress, Saeedi, & Brac-de-la Perriere, 2014; Linao, 2016; Timekeeper, 2017). In fact researchers have been working with Augmented Displays for more than 20 years (Bhorkar, 2017; Mann, 2001, 2004, 2013; Mann & Fung, 2002; Starner et al., 1997; Tang, Aimone, Fung, Marjan, & Mann, 2002). 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 (Ackerman, 2013; Hayes, 2016; Mann, 2013). Glass, as it became known, was considered to be a game changing technology for a few years across a wide range of industry sectors (Elise, 2014; Fox & Felkey, 2013; Muensterer, Lacher, Zoeller, Bronstein, & Kübler, 2014; Schreinemacher, Graafland, & Schijven, 2014). But ultimately, the product was withdrawn (Cave, 2015).

A key challenge with AR/VR headsets is that, lacking a physical keyboard they do not provide an intuitive mean of user authentication. The weak authentication available in Glass (Ching & Singh, 2016; Hayes, 2016; Yadav, Ionascu, Ongole, Roy, & Memon, 2015) subsequently led to various attempts to refine and improve on the basic authentication of the headset (Chan, Halevi, & Memon, 2015; Chauhan, Asghar, Kâafar, & Mahanti, 2016; Peng et al., 2017). Ultimately, the device authentication was simply not adequate and led, in part, to its withdrawal from the market.

This leads us to consider how the next generation of wearable AR/VR vision systems might implement a more seamless and intuitive authentication mechanism without sacrificing security and robustness. The implementation of a face recognition system is not practical, as the form-factor of an AR/VR head-set does not allow to capture a full facial image. However, with the reduction in size and cost of multi-cameras systems on mobile devices it is now practical to consider that rear-facing (i.e. user-facing) camera systems can be incorporated into such headsets. One important driver for rear-facing cameras is the use of eye-tracking to dynamically determine the wearer's point of gaze (PoG) which is important for accurate AR/VR rendering (Cognard, Goncharov, Devaney, Dainty, & Corcoran, 2018; Rompapas et al., 2017).

Iris authentication is a proven and reliable biometric trait with high distinctiveness, permanence and performance (Prabhakar, Pankanti, & Jain, 2003). The use of iris recognition on consumer devices is explored across multiple works (Corcoran, Bigioi, & Thavalengal, 2015; Shejin & Corcoran, 2016; Thavalengal, Andorko, Drimbarean, Bigioi, & Corcoran, 2015; Thavalengal, Bigioi, & Corcoran, 2015b) and the importance of accurate iris segmentation, particularly in consumer imaging devices, is identified as a key challenge (Bazrafkan, Thavalengal, & Corcoran, 2018; Thavalengal, Bigioi, & Corcoran, 2016). In the iris authentication workflow, failed segmentations represent the single largest source of error (Erbilek, Da Costa-Abreu, & Fairhurst, 2012; Hofbauer, Alonso-Fernandez, Bigun, & Uhl, 2016; Proença & Alexandre, 2010). 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 (Hammal, Massot, Bedoya, & Caplier, 2005). Eye-gaze as mentioned is a key element of various user-interface modalities for wearable AR/VR displays.

1.1. Background to the problem

Data augmentation techniques and their effectiveness in computer vision have been explored in multiple research works.

In Shijie et al. (2017) and Taylor and Nitschke (2017) the effect of several data augmentation techniques (GAN/WGAN, Flipping, Cropping, Shifting, PCA jittering, Colour jittering, Noise, Rotation)

is investigated, in improving the performance of a CNN in the image classification task. Furthermore, researchers have studied ways to implement and automate different data augmentation techniques (beyond the traditional techniques) that can boost a CNN's performance.

In Perez and Wang (2017) the authors present an approach called neural augmentation, which allows a neural network to learn augmentations that will improve the accuracy of a CNN classifier. Similarly in Cubuk, Zoph, Mane, Vasudevan, and Le (2018) where their Autoaugment method searches automatically for improved data augmentation policies in order for a neural network to yield higher performance in the validation set.

Additionally in Lemley, Bazrafkan, and Corcoran (2017), the Smart Augmentation method learns suitable augmentations during the training of the neural network. The results indicate that this method could be applied for a range of tasks and that by implementing this augmentation method a small network achieved better results than those obtained by a much larger network. Finally the effectiveness of data augmentation is not applicable only to computer vision problems but also in audio (Salamon & Bello, 2017; Schlüter & Grill, 2015). It is worthwhile to mention that these methods are solely designed for classification problems and data augmentation for regression problems are widely applied by prior knowledge.

As new technologies arise, deep neural networks can provide solutions to the consequently new challenges that appear, but the data collection related to the new problems and their annotation remains the bottleneck of the process. Data augmentation is commonly used to boost the performance and the generalization of the CNNs but with specialized techniques can be utilized to eliminate the expensive procedure of the data collection–annotation. By designing specific data augmentation techniques, it is possible to simulate the characteristics and features related to the new challenges and inject them in relation to the problem datasets, so that a CNN can be efficiently train for the new task without going through a data acquisition and annotation procedure.

In this case the new challenges arise from the next generation of wearable AR/VR vision systems which as mentioned in the introduction will have to implement a more seamless and intuitive authentication mechanism that is available on today's mobile devices. Therefore, the next generation of AR/VR headsets might incorporate a user-facing camera installed below the eye level, used for iris authentication or eye-gaze tracking. In such a case, off-axis iris images can be readily obtained and provide a suitable biometric for user authentication.

Note that a key challenge for accurate iris recognition is to accurately segment the iris region (Bazrafkan et al., 2018; Jillela & Ross, 2013). Given that camera locations for AR/VR devices must be mounted off-axis, often with an oblique perspective and close proximity to the observed eye-region the segmentation process for such off-axis iris regions becomes even more critical as errors at the segmentation stage are propagated to the feature extraction and pattern matching stages of the authentication workflow (Bazrafkan et al., 2018; Hofbauer et al., 2016; Hofbauer, Alonso-Fernandez, Wild, Bigun, & Uhl, 2014). While there are past studies on off-axis iris and its effects on recognition rates, the problem of near-eye iris segmentation is a new problem arising from the introduction of emerging AR/VR headset technology into consumer devices and a main obstacle for providing a solution, is the unavailability of a dataset with iris samples captured from an user-facing camera on AR/VR devices, but data augmentation techniques can be utilized to eliminate this obstacle.

Currently, the majority of existing iris recognition systems follow the authentication workflow as (i) image acquisition: an eye image is acquired using a camera, (ii) iris segmentation: eye/iris region is located in this image followed by isolating the region

representing iris. (iii) iris normalization (iv) Feature extraction: relevant features which represent the uniqueness of the iris pattern is extracted from the iris region and (v) similarity of the two iris representations is evaluated by pattern matching techniques. The described workflow is illustrated in Fig. 1, highlighting the focus of this work, which is on the second step of the iris recognition workflow, the iris segmentation. Specifically, specialized data augmentation techniques are designed to simulate the iris images as represented when captured from a user-facing camera on an AR/VR device along with their ground truth segmentation map and thus eliminating the expensive process of data acquisition and annotation. In continuance a low-complexity network is designed and trained with the augmented samples to successfully segment the off-axis close proximity iris images. Although this work is focused solely on the use of deep neural networks for the off-axis iris segmentation task, it could be widely applicable for other off-axis region segmentation problems.

1.2. Related literature

The significant success of deep neural networks at vision-oriented tasks has enabled exceptional advancement in semantic segmentation. An instance of that is the DeepLab method (Chen, Papandreou, Kokkinos, Murphy, & Yuille, 2014), where instead of using deconvolution, they proposed Atrous ('Holes') convolution. The proposed method is combined with fully connected conditional random fields (CRF) and is able to produce semantically accurate predictions and detailed segmentation maps efficiently. In a follow-up publication (Chen, Papandreou, Kokkinos, Murphy and Yuille, 2018), the same team proposed an atrous spatial pyramid pooling (ASPP) module, consisting of multiple parallel atrous convolutional layers with different sampling rates to improve the segmentation of objects. The DeepLab team consistently proposes methods for improving the segmentation standards (Chen, Papandreou, Schroff, & Adam, 2017; Chen, Zhu, Papandreou, Schroff and Adam, 2018). In another semantic segmentation approach (Jiang, Yuan, & Wang, 2018), the authors develop an adaptive-depth neural network to obtain the coarse semantic segmentation results. At the same time, the contour information is provided by a contour-aware network. Both the coarse semantic information and contour information are modelled in the same way and combining this information the semantic labels are given through global inference based on CRF. Furthermore, semantic segmentation methods could provide a solution to road detection, a challenging problem in autonomous driving. In Wang, Gao, and Yuan (2018), a Siamese neural network is developed which accepts an RGB image, the semantic contour and the location prior, at the same time, for the road detection task and the results demonstrate that the network is able to learn discriminative features of road boundaries and location prior. More detailed information for semantic segmentation techniques using deep learning can be found in Garcia-Garcia, Orts-Escalano, Oprea, Villena-Martinez, and Garcia-Rodriguez (2017) and Lateef and Ruichek (2019).

As the work is focused on the iris segmentation task, a quick overview of similar works in the literature is outlined below.

1.2.1. Frontal iris segmentation

The number of methods in the literature regarding iris segmentation shows that this topic has been thoroughly studied but remains an active area of research. When referring to iris segmentation algorithms, a good starting point is two highly cited works in the literature: Daugman (2009) and Wildes (1997). In the iris matching algorithms developed in these research papers iris segmentation is achieved by fitting a circular contour to the iris and pupil. These two methods differ mostly in the way

they define the circular boundaries on the image information. Daugman's integrodifferential operator searches the entire image pixel by pixel to find the best circular path for the iris and pupil boundaries. While Wildes, in order to fit the circular contour, combines an edge detector and Hough transform.

In continuance approaches were implemented in an attempt to speed up the process. For example, Liu, Yuan, Zhu, and Cui (2003), uses a Canny edge detector with a Hough transform to provide a fast localization of the iris edges with the assumption that the iris texture is located between two homocentric circles. Several other methods were developed based on Wildes's and Daugman's implementations such as: Huang's (Huang, Luo, & Chen, 2002), Khan's (Khan et al., 2011), He's and Shi's (He & Shi, 2006), Lili's and Mei's (Lili & Mei, 2005).

As noted, the aforementioned methods assume the circularity of the iris outer boundary and pupil boundaries. However, Daugman in his follow up work (Daugman, 2007) shows that a non-circularity applies to the iris and pupil contour which when defined precisely has a significant influence on recognition performance. Therefore, adopts an active contour or snake model to segment the iris. Furthermore, Shah and Ross (2009) implemented a geodesic active contour to capture the iris texture and experimental result on non-ideal iris images designates the effectiveness of this method. Koh, Govindaraju, and Chaudhary (2010) similarly implemented an active contour model which was combined with the Hough transform for iris localization. In another approach, Broussard and Ives (2009), used a feature saliency algorithm to identify the measurements that could define the iris boundary. The selected measurements are fed to a shallow artificial neural network in order to accurately predict the outer iris boundary. A detailed overview of the iris segmentation literature can be found in Bowyer, Hollingsworth, and Flynn (2008, 2013), as well as in Jan (2017) where approaches for segmenting non-ideal iris images are reviewed.

1.2.2. Off-axis iris segmentation

A subsection of non-ideal iris images includes the off-axis iris images. Localizing the iris in this type of images has always been a challenge for researchers. In Dorairaj, Schmid, and Fahmy (2005), Dorairaj assumes that a rough estimation of the angle rotation is available in order to deal with the off-axis iris problem. Two different objective functions are used to refine the estimate. When two images are available from the same iris class, the "ideal" and off-axis iris image, the Hamming distance between the ICA coefficients of the two images is calculated. In the case that only the off-axis image is available, Daugman's integro-differential operator is used. A projective transformation is applied to rotate the off-axis image into a frontal view image once the angle is estimated. In the next step, the image is enhanced and segmented with the integro-differential operator. In another approach, Li in Li (2006) first fits an ellipse to the pupil boundaries. After that based on the information that has been retrieved from the ellipse fitting, rotation and scaling are applied to the image, to restore the straight position of the ellipse and the circularity of the pupil. The segmentation of iris is then operated by Daugman's like algorithms. A similar approach can be found in Abhyankar, Hornak, and Schuckers (2005) where the use of projective and affine transformation is explored in order to bring the off-axis iris images and match them with frontal iris images. This approach comes with some serious downsides, such as the blurring of the iris outer boundaries and the fact that a prior knowledge of the angle is required for the transformation. Finally, in Abhyankar and Schuckers (2006) the use of active shape models to retrieve the elliptical boundaries of the off-axis iris is investigated.

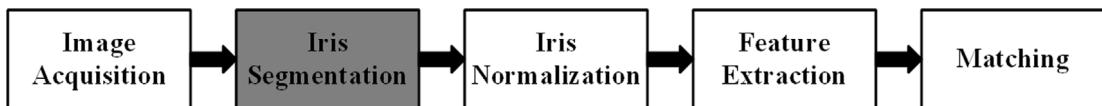


Fig. 1. Iris authentication workflow. In practical implementation the bulk of authentications errors are due to incorrect segmentations (Erbilek et al., 2012; Hofbauer et al., 2016; Proen  a & Alexandre, 2010).

1.2.3. Deep learning approaches for iris segmentation

Liu, in Liu et al. (2016) proposed two CNN approaches to segment noisy iris images acquired under unconstrained conditions. In the first approach called hierarchical convolutional neural networks (HCNNs), three patches taken from different scales of the same image are used as input. The HCNN consists of three similar blocks, a combination of convolutional and pooling layers that are merged together into a fully connected layer. In the second approach, 31 convolutional layers and 6 pooling layers are used to compose the multi-scale fully convolutional network (MFCNs). Both models are end-to-end, with no requirement for pre- or post-processing of the image. Arsalan et al. (2017), introduced a two-stage iris segmentation method. The first stage includes a pre-processing of the image and the use of a modified Hough Transform to identify the region of interest (ROI). In the second stage, a mask of $[21 \times 21]$ pixels, based on the ROI defined in the previous stage, is fed to a pre-trained VGG-face model which classifies the pixels as iris or non-iris. In a follow up work which is focused on segmenting low quality iris images, Arsalan in Arsalan et al. (2018), proposed a densely connected fully convolutional network (IrisDenseNet), consisting of two main components: a densely connected encoder and a SegNet decoder. In a similar work, Bazrafkan in Bazrafkan et al. (2018), presented a network design focused on segmenting iris of inferior quality. Four different end-to-end fully convolutional networks are merged into a single model using a method known as Semi Parallel Deep Neural Networks (SPDNN). In this way, the final model benefits from each of the four distinct network designs. Furthermore from a more medical aspect in Lakra, Tripathi, Keshari, Vatsa, and Singh (2018), utilizes a DenseNet-121 which has four convolution blocks to be able to segment iris with cataract or post cataract surgery. Finally, since the existence of a large labelled dataset is a prerequisite in order to implement a convolutional neural network approach, Jalilian in Jalilian, Uhl, and Kwitt (2017) to overcome this obstacle, introduced a domain adaption method so that a CNN for iris segmentation could be trained with a limited data.

1.3. Contributions

The focus of this work is to improve the segmentation of off-axis iris images originating from the unconstrained conditions of a user-facing camera on wearable AR/VR device.

The model proposed is an end to end deep neural network which accepts an off-axis eye-region image and generates the corresponding binary segmentation map for the iris region as output. Performance evaluation of the proposed model shows advantages over recent iris segmentation techniques in the literature which together with its simple, yet efficient design makes it well-suited for deployment in wearable AR/VR devices.

Three noteworthy contributions are presented in this work.

1. Specialized data augmentation methods that generate distorted iris images of size and quality typical of the user-facing camera employed on today's wearable AR/VR headsets. These are derived from a high-quality iris dataset together with a corresponding ground truth.

2. By utilizing the data augmentation techniques and producing a large number of representative data, it enabled us to propose an improved low complexity neural network design for the off-axis iris segmentation task with reduced memory and computational requirements in comparison with other deep learning state-of-the-art iris segmentation techniques while achieving equivalent performance.
3. A thorough evaluation of the proposed segmentation model is presented on several well-known public iris datasets. The presented method is compared with state-of-the-art iris segmentation techniques.

1.4. Foundation methods

1.4.1. Data augmentation

Despite the large amount of data available today, there are many problems where data collection and annotation poses challenges and thus only small datasets are publicly available. In some cases, the data contains sensitive information such as in medical applications or due to privacy/legislations reasons, data is not easily accessible. Also, with the technology rapidly growing, new problems arise frequently and in many cases it takes a while before a proper dataset is built and made publicly available. The problem investigated in this work is a clear example of the later situation. Therefore, data augmentation can be utilized not only to increase the performance of a CNN, but also to overcome the non-availability of data due to the aforementioned reasons. Applying the appropriate augmentation techniques to available datasets allows to investigate a problem where there are currently no existing public datasets and where the collection of a large training dataset poses significant challenges. Data augmentation can simulate the features and characteristics of such problems without the time, expense and data-collection and annotation challenges associated with building a dataset of many thousands of individual subjects.

In regard to our problem the main focus of the proposed augmentation techniques, is to simulate off-axis iris images as captured by a user-facing camera on AR/VR device, as to the best of our knowledge such dataset is not available. As mentioned, a possible location of the user-facing camera utilized for iris recognition and eye-gaze is below the eye. In that case the iris samples obtained will be off-axis in the horizontal and the vertical plane. Main characteristics of the iris images taken with a head-mounted device are their elliptical shape along with the fact that are not centred, in contrast with frontal iris images where the iris is most of the time centred and a more circular shape is obtained. Therefore, the augmentation techniques in this work are focused in achieving the described representation. In addition, a secondary goal of our augmentation is to introduce effects of images when captured in real-life, where the samples are not obtained in constrained conditions and a lower quality is reported. These augmentation techniques are focused on reducing the contrast between the iris and the pupil as well as adding noise to the samples.

1.4.2. Network design

In this work a low complexity network, targeted for deployment in embedded devices is designed and trained to generate the segmentation map for low quality off-axis iris images.

In order to achieve high performance results, when a network is designed, large structures with high capacity are favoured. That is translated into CNNs containing millions of parameters, which to be used require large memory and high operation cost. Therefore, executing deep CNNs requires significant hardware resources which is a limited specification in many computational platforms.

The number of parameters in the proposed network is significantly lower compared to the parameters of other deep learning approaches designed for the iris segmentation task. Thus, making the proposed network faster and with reduced memory requirements, while attaining high performance results in producing the segmentation map for off-axis iris images of low quality as represented when captured by a user-facing camera on AR/VR headset and therefore well-suited for deployment in such devices.

The rest of the paper is arranged as follows: In Section 2, the datasets used are presented along with a detailed description of the augmentation techniques. In Section 3, the network design and training are explained. In Section 4, the results are illustrated and in Section 5, the numerical evaluation of the proposed method and its comparisons with state-of-the-art segmentation techniques are presented.

Finally, it should be remarked that preliminary results from this research were first presented in Varkarakis, Bazrafkan, and Corcoran (2018). This article builds on that earlier work with more detailed and extensive experimental verifications, exhaustive description of the augmentation techniques and direct comparison on off-axis and frontal iris samples with state-of-the-art iris segmentation techniques.

2. Datasets and augmentation methodology overview

In this work, three datasets are utilized. CASIA Thousand (“CASIA Iris Image Database”, 2019) and Bath800 (Rakshit, 2007) are used during the training and testing stages. UBIRIS v2 (Proenca, Filipe, Santos, Oliveira, & Alexandre, 2010) is used for tuning and testing. Two types of augmentation methods are described below. The first type is concentrated on adding real-world condition effects to the iris images, while the second is focused on augmenting the images so that they represent off-axis iris images. The combination of these two types of augmentation methods results in iris images as captured by an user-facing camera on AR/VR device. Below, the datasets used are presented along with the production of their ground truth, and finally, the augmentation techniques are explained.

2.1. Datasets

CASIA-Iris-Thousand is a subset of CASIA-Iris V4 dataset. This subset contains 20 000 iris images from 1000 subjects. The iris images are constrained, high quality and high contrast. Bath800 dataset is made of 31 997 images taken from 800 individuals. The samples similarly to the CASIA Thousand are of high quality and high contrast. Both datasets consist of Near InfraRed (NIR) samples. Finally, UBIRIS v2 dataset includes 11 102 iris images from 261 subjects, captured in visible wavelength. The samples are of low-quality as they are taken under unconstrained conditions. More detailed description of CASIA Thousand, Bath800 and UBIRIS v2 can be found in Bazrafkan et al. (2018). Samples from the datasets used in this work are shown in Figs. 2–4.

2.2. Ground truth

Bath800 and CASIA Thousand are not provided with the segmentation ground truth. However, these datasets as mentioned above contain images of high quality, high contrast and are captured under constrained conditions. In this work, the binary iris map for these datasets is produced using the commercial iris segmentation solution MIRLIN (“MIRLIN”, 2019). The obtained segmentation map is considered in this work as the ground truth. The selection of the segmentation algorithm is based on the availability as well as its performance on large-scale iris evaluations (Quinn, Grother, Ngan, & Matey, 2013). The same segmentation solution was also adopted in Bazrafkan et al. (2018). The low-resolution segmentations for Bath800 and CASIA Thousand are publicly available.¹

Regarding UBIRIS v2, the manual segmentation generated by WaveLab² (WaveLab, 2019), available in IRISSEG-EP dataset (Hofbauer et al., 2014), is used. The manual segmentation map is not available for all the samples of the dataset. Segmentation of only 2250 images from 50 individuals is provided and therefore only these are used in this work. Segmentations examples derived from these datasets are shown in Figs. 5–7.

2.3. Data augmentation

In order to accurately train a deep neural network, a large number of labelled training samples are required. These samples should correctly characterize the imaging problem so that it enables the deep learning process to train an accurate model. To the extent of our knowledge, there is not available a dataset with iris samples captured from a user-facing camera on AR/VR device. Even if such datasets were available it would require an accurately marked segmentation ground truth – a task which poses new problems over more conventional frontal iris images. Thus, in order to obtain a large number of labelled samples to enable the training of a DNN for AR/VR iris segmentation task, some specialized augmentations of existing datasets are required. To find the best augmentations for the iris images, precise observations have been made on iris images obtained by a user-facing camera on head-mounted displays.

The augmentation techniques are divided into two categories. The first category of augmentation techniques is focused on representing real-life scenarios where low-quality images are obtained. Based on research that has been done in Thavalengal et al. (2015b) and Thavalengal, Bigioi, and Corcoran (2015a), the difference between high-quality constrained iris images and wild ones is linked to contrast, blurring, and shadows. Consequently, to simulate the effects of real-world conditions in iris images the contrast is changed, motion blurring and shadows are added to the images. The augmentation techniques used to deteriorate the image quality and simulate unconstrained conditions are derived from Bazrafkan et al. (2018). The objective of the second category's augmentation techniques is to simulate the representation of iris images as captured by a user-facing camera on an AR/VR device. This representation includes off-axis iris images mainly of elliptical shape and not centred in the image.

The augmentation techniques are detailed in Sections 2.3.1 and 2.3.2. The workflow that is followed for the augmentation of the datasets is described in Section 2.3.3. In this work all the samples are resized to [120 × 160] using bilinear interpolation. Smaller resolution samples are preferred rather than larger ones as it accelerates the training of the deep neural network.

¹ <https://Goo.gl/JVksyG>.

² <http://www.wavelab.at/sources/Hofbauer14b>.



Fig. 2. Eye socket samples from Bath800 dataset.

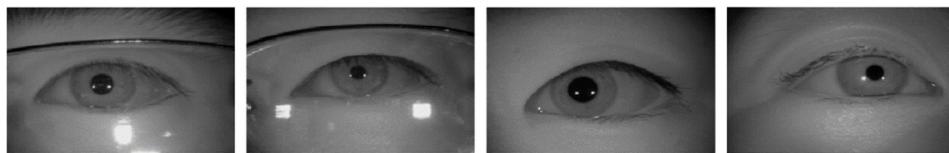


Fig. 3. Eye socket samples from CASIA Thousand dataset.

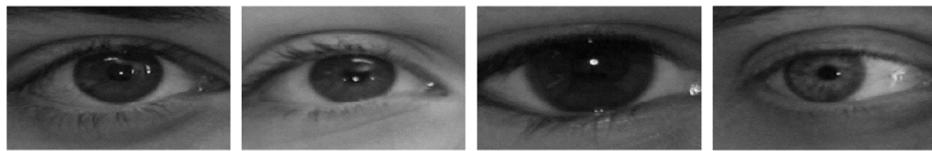


Fig. 4. Eye socket samples from UBIRIS v2 dataset.



Fig. 5. Bath800 automatic segmentation results.

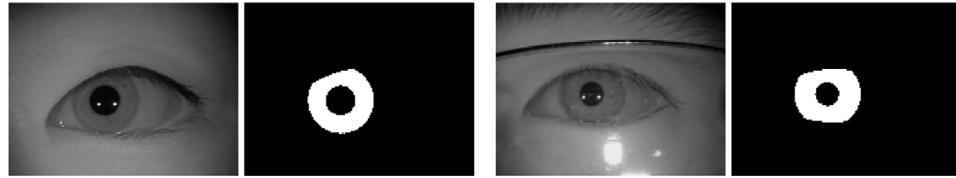


Fig. 6. CASIA Thousand automatic segmentation results.



Fig. 7. UBIRIS v2 manual segmentation results.

2.3.1. Data augmentation: Simulating unconstrained conditions

The first type of augmentation techniques is applied to ensure that the samples used to train the network represent real-life scenarios. The distribution of the input data plays a vital role in what the network learns and how it will behave during the testing stage but also in unconstrained situations. As mentioned earlier, to simulate real-life captured iris images of low quality, the contrast of the samples is changed, blurring and shadows are added to the samples with the following augmentation techniques. The techniques mentioned below are derived from [Bazrafkan et al.](#)

(2018) and used with slight changes. The original code of these augmentation techniques is available.³

2.3.1.1. Augmentation 1: Image contrast. The iris images captured by an AR/VR device in real-world conditions compared to the high-quality, high-resolution NIR iris images acquired in constrained conditions have significant differences. The differences are with regard to the amount of contrast inside and outside

³ https://github.com/C3Imaging/Deep-Learning-Techniques/blob/Iris_SegNet/DAugmentation/DBaug.m.

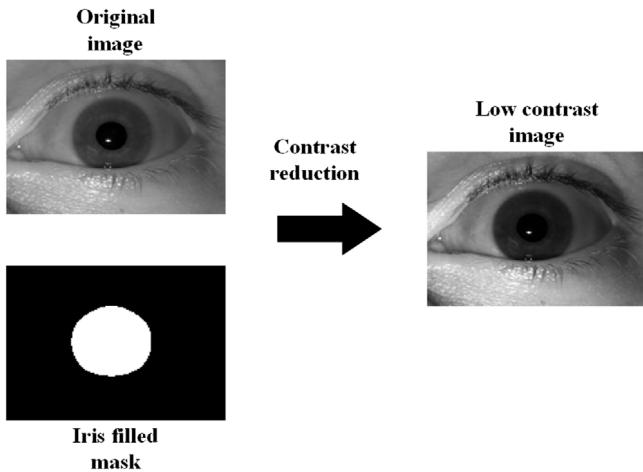


Fig. 8. For inside the iris region, the contrast is reduced, and the region is getting darker. The outside of iris is altered by decreasing the contrast.

the iris region as in unconstrained scenarios the samples are suffering from low contrast. Another difference noted is the intensity properties of the low-quality samples inside and outside the iris region. The region inside the iris is darker than the same region in high-quality samples. For the outside region of the iris the level of brightness cannot be categorized as it could differ from overexposed and strongly bright till very dark. To bring these properties to high-quality images, the contrast inside and outside the iris region is modified separately. This is achieved with the use of histogram mapping. The following histogram mapping equations are used to reduce the contrast of the iris images. Eq. (1) is used for the region outside the iris and (2) is used for the region inside the iris.

$$y_{out} = \text{norm}(\tanh\left(3 * \left(\frac{x}{255} - 0.5\right)\right) + \mathcal{U}(-0.2, 0.3)) * 255 \quad (1)$$

$$y_{in} = \text{norm}(\tanh\left(3 * \left(\frac{x}{255} - 0.45\right)\right) - \mathcal{U}(0, 0.2)) * 255 \quad (2)$$

where x is the input intensity in the range $[0, 255]$, y is the output intensity in the same range, $\mathcal{U}(a, b)$ is the Uniform distribution between a and b , and the norm function normalize the output between 0 and 1. As mentioned above the outside and inside regions of iris suffer from low contrast, but the brightness differs. For the region outside of the iris, the histogram mapping with Eq. (1) can result in bright, dark, or normally exposed low contrast outputs. For the region inside the iris, where Eq. (2) is used, the contrast is reduced while the brightness of the iris region is reduced as well. Different equations are used to reduce the contrast in the inside and outside regions of the iris so that variety is obtained. An example of this step is shown in Fig. 8.

2.3.1.2. Augmentation 2: Motion blur. Wearing AR/VR devices, head movements are inevitable. These movements can cause motion blur. Therefore, to mimic these situations and train the model in order to be efficient in these cases, motion blurring has to be introduced to the training images. In order to include this effect, the image is passed through a motion blur filter, applying the linear camera motion by $\mathcal{U}(3, 7)$ pixels in the direction $\mathcal{U}(-\pi, \pi)$, where $\mathcal{U}(a, b)$ is the Uniform distribution between a and b . The low contrast image after applying motion blur is shown in Fig. 9.

2.3.1.3. Augmentation 3: Shadowing. In unconstrained conditions, the illuminations scenarios vary. One main effect produced by different illumination directions is shadows. In order to add this effect, the iris images were multiplied with the following shadow

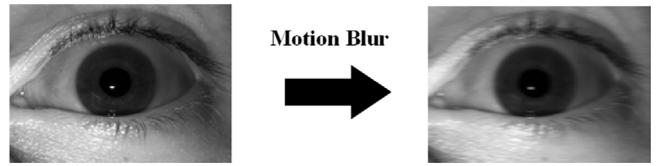


Fig. 9. Applying motion blur in a random direction to the low contrast image.



Fig. 10. Shadowing applied to low contrast blurred image.

function:

$$y = \text{norm}(\tanh(2 * \text{randSign} * (x - 0.5 + \mathcal{U}(-0.3, 0.3))) + \mathcal{U}(0, 0.1)) \quad (3)$$

where x is the dummy variable for image column number and y is the coefficient for intensity, $\mathcal{U}(a, b)$ is the Uniform distribution between a and b , the norm function normalizes the output between 0 and 1, and the *randSign* generates a random coefficient in the set $\{-1, 1\}$ which determines the direction of the shadow. The final image after applying shadowing is given in Fig. 10.

The segmentation map for these augmented samples is the same as the original segmentation ground truth as the structure and position of the iris remain unchanged.

More detailed information regarding the augmentation techniques simulating unconstrained conditions can be found in Bazrafkan et al. (2018) from where are originated.

2.3.2. Off-axis, near-perspective iris data augmentation

The second category contains two augmentation techniques, which their goal as mentioned previously is to generate iris images as they appear when acquired by a user-facing camera on AR/VR device. As noted in the introduction, a possible location of the camera used for obtaining an iris image that is to be used in iris recognition or eye-gaze is below the eye. Therefore, the iris images captured are off-axis in both horizontal and vertical planes. The augmentation techniques described below are specialized to produce such off-axis iris images. Furthermore, the combination of the two augmentation techniques and the multiple ways and different volume that each technique can be applied to an image, it allows to represent the iris regions as captured from a user-facing camera on AR/VR device from many varying perspectives, as multiple AR/VR headsets exist and each will locate the camera in a different position. This is a desired goal as it facilitates the training of generic DNN that is able to segment the off-axis iris region from user-facing cameras installed on AR/VR devices and being invariant of the camera's different set-ups. The code for these augmentation techniques is also available.⁴

2.3.2.1. Augmentation 4: Spatial stretching/contracting. The iris images when captured from an AR/VR, are characterized as distorted and with an elliptical shape. In addition, the iris is not at the centre of the image as usual. In order to generate iris images with these properties, the samples are warped by applying a spatial stretching/contracting to the iris images. The stretching is linearly applied to the images. The stretching is achieved by

⁴ https://github.com/C3Imaging/Deep-Learning-Techniques/tree/Off_axis_Iris.

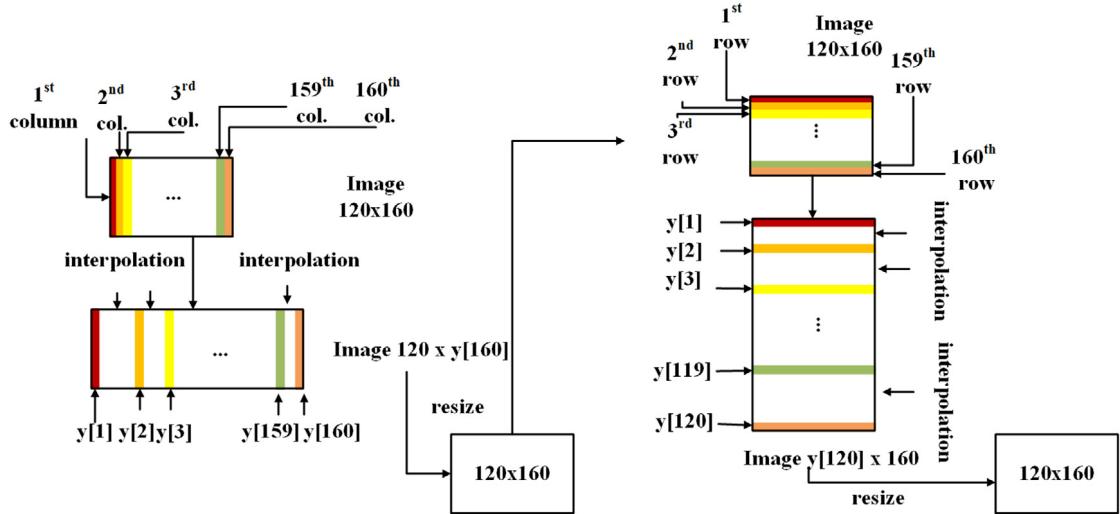


Fig. 11. Workflow of spatial stretching/contracting, illustrating the mapping of columns/rows to a new position based on $y[j]$.

mapping every column/row of the image to a new position given from $y[j]$ as shown in Fig. 11.

The equations below illustrate how $y[j]$ is calculated for columns/rows:

$$\lambda = \mathcal{U}(2, 17) \quad (4)$$

$$k[i] = \frac{\left(\frac{1}{\lambda}\right) - \lambda}{s - 1} * t[i] + \lambda, i \in [1, s] \quad (5)$$

$$k[i] = \frac{\lambda - 1}{s - 1} * t[i] + \lambda, i \in [1, s] \quad (6)$$

$$a = a + k[j], j \in [2, s] \quad (7)$$

$$y[j] = \text{Round}(a * 5), j \in [2, s], y[1] = 1 \quad (8)$$

where s is the length of the columns or rows of the original image depending on where the distortion is applied, $t[i]$ is a vector which includes all the integer values $[0, s - 1]$ in ascending order and $\mathcal{U}(a, b)$ is the Uniform distribution between a and b .

The first column/row of the original image is mapped on the first column/row of the stretched image. The following columns/rows are then mapped in the position determined by the value of $y[j]$. Depending on which is the desired direction for stretching the image, Eq. (5) or (6) is used in calculating $y[j]$.

The combination of (5) and (6) makes it possible to stretch the images in four main directions. If (5) is used for mapping the columns and the rows, the image will be stretched in the right and down direction. In case (6) is used for both columns and rows, the image is stretched at left and up. Using (5) when mapping the columns and the (6) when mapping the rows of the image, will result in stretching the image to the right and up direction. Finally mapping the columns using (6) and the rows using (5), the image will be stretched to the left and down direction. Each direction has the same probability of being selected when the image is stretched. For each distortion and each direction, the amount of stretching applied to the images differs on every occasion as well as the volume of the stretching applied to the columns and the rows of the image is different, so that variation is injected to the augmented dataset. The stretching is applied at first to the columns of the image. The void spaces that are created, are interpolated with a weighted nearest neighbour method, which is explained by the following equations:

$$c[i] = \frac{\frac{f(y[j])}{i-y[j]} + \frac{f(y[j+1])}{y[j+1]-i}}{\frac{1}{i-y[j]} + \frac{1}{y[j+1]-i}}, j \in [1, 160], i \in (y[j], y[j+1]) \quad (9)$$

where $f(x)$ is a function that returns the values of the x th column/row, $c[i]$ represents the values of the i th column/row of the stretched image. The values of $y[j]$ and $y[j + 1]$ are the positions where the columns or rows of the stretched image have values and the columns/rows that need to be interpolated are located between these two positions. Finally, the image is contracted as the image is resized to the original resolution $[120 \times 160]$ using bicubic interpolation. The same process is then applied to the rows of the image. The same workflow is used for the ground truth segmentation map in order to obtain the segmentation map for the augmented sample.

The described workflow for the spatial stretching/contracting of an image is illustrated in Fig. 12. Applying spatial stretching/contracting results in an iris region that is not located in the centre of the image and with non-circular iris-pupil structures, as shown in Fig. 13 which is a usual case in iris images acquired from a user-facing camera on AR/VR headsets.

2.3.2.2. Augmentation 5: Image tilting. A possible location of the camera used for capturing the iris images, as mentioned in the introduction, will be below the eye. Therefore, the iris images should be representing samples which when captured, the camera is positioned below the eye level. To achieve that effect and also give an elliptical shape to the iris, in this second augmentation technique the samples are tilted in two directions: up and left, up and right.

A projective transformation is applied to the images. This transformation maps the top vertices of the image to a new pair of points as illustrated in Fig. 14. The values from Fig. 14, a , b , c , and d are randomly generated between a range of values, so the image is tilted in the desired direction with variation. When the image is tilted up and left the values of a , b are in $\mathcal{U}(0.15, 0.45)$, c in $\mathcal{U}(0.9, 1)$ and d in $\mathcal{U}(0, 0.1)$. When the image is tilted up and right, the values of a , b are in $\mathcal{U}(0, 0.1)$, c is in $\mathcal{U}(0.55, 1)$ and d is in $\mathcal{U}(0.15, 0.45)$, where $\mathcal{U}(a, b)$ represents the Uniform distribution between a and b . During this transformation as the image shrinks the interpolation used is the nearest-neighbour. The probability of the images being tilted in a direction between the two options (up and left/up and right) is the same.

As shown in Fig. 14, when the transformation is applied, while mapping the top vertices, to the a , b , c and d points, the image is compressed at the boundaries. Since the resolution of the image has to stay unchanged, the void spaces around boundaries should be filled to avoid sharp edges in the image. Since the void spaces are at the boundaries of the image, there is not a direct way

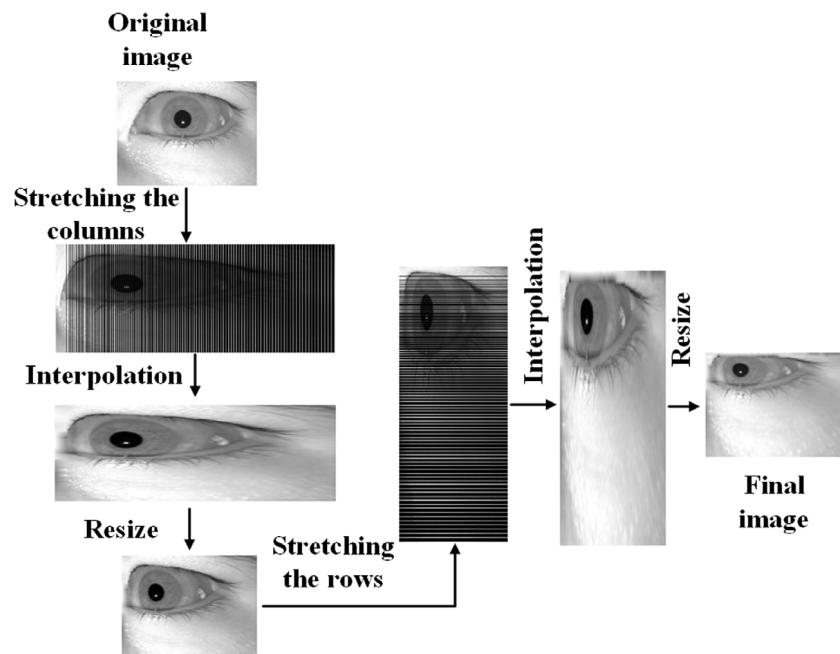


Fig. 12. Workflow of spatial stretching/contracting. For this transformation Eqs. (6) was used to map the columns and the rows of image and direct the image in up and left direction.

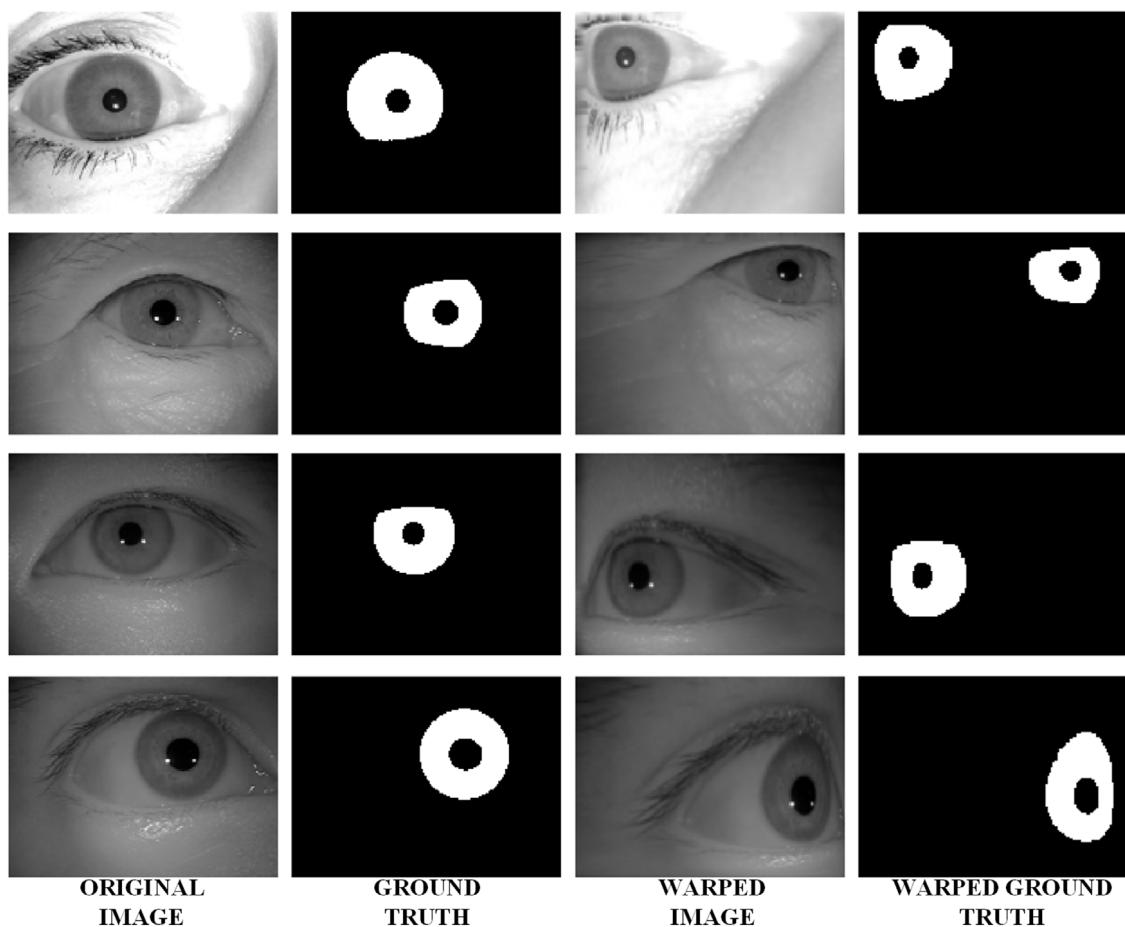


Fig. 13. Spatial stretched/contracted (warped) samples and their corresponding segmentation map.

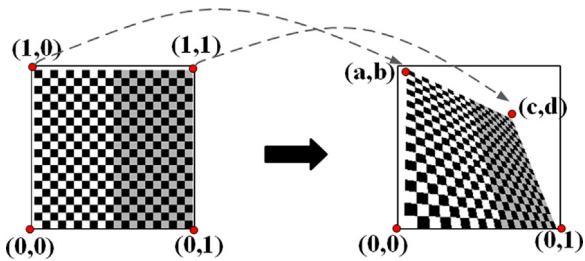


Fig. 14. Tilt transformation.

to apply interpolation. Therefore, the value from the edges of the tilted image is extended for each column up to the image boundary. The same process is applied to the image rows. After this process has been applied in both columns and rows, the average value is assigned to the void spaces. Finally, in order to smooth the interpolated areas of the image, a gaussian 3×3 filter with standard deviation σ equal to 2 is applied to this region.

The described workflow for tilting an image is shown in Fig. 15, and in Fig. 16 samples are shown where the tilting transformation is applied. To obtain the segmentation map for the augmented samples, the same workflow is applied to the segmentation ground truth with the only difference being that the void spaces created are filled with black.

2.3.3. Dataset preparation

2.3.3.1. Workflow of combining the augmentations techniques. The augmentation techniques are combined in various ways so that the dataset represents a generalized and realistic scenario and as a result the trained model can be robust and perform well in all the different conditions that one can encounter with iris images acquired from a user-facing camera on AR/VR device.

The augmentation techniques are mixed in three ways. Samples are augmented by only using the methods for simulating the off-axis, near-perspective iris images. At first, the spatial stretching/contracting transformation is applied to an image with 50% probability. In the next step, the tilting transformation is applied to the rest of the samples that the first transformation was not applied to. In addition to these, for an image that spatial stretching/contracting is applied to in the first step, there is a 50% probability that tilting is applied afterwards. In the second way, the samples are augmented using only the methods for simulating unconstrained conditions. At first, the contrast of all the iris images is modified as explained. Afterwards, all the images are passed through the motion filter, and finally, the technique used to introduce shadows is applied to the image. The probability that shadows are added to an image is 50%. Thirdly, the techniques from the two augmentation categories are combined. Initially, the techniques simulating the off-axis, near perspective iris images are applied to an image based on the augmentation workflow described above. Later the same image is processed using the techniques simulating unconstrained conditions, including contrast reduction, motion blurring, and shadowing. In Fig. 17, the workflow explained is illustrated.

The augmented samples simulate iris images captured using a user-facing camera on AR/VR device, frontal iris images affected by unconstrained conditions and AR/VR images affected by unconstrained conditions. Bath800 and CASIA Thousand are augmented with all three combinations of the augmentation techniques described. UBIRIS v2 was augmented only by using the augmentation techniques simulating iris images as represented by an AR/VR device. The samples of this dataset as mentioned previously are captured in unconstrained conditions, and therefore it will be redundant to make use of the augmentation techniques that simulate real-world conditions as they already exist in the dataset of UBIRIS v2.

2.3.3.2. Dataset analysis. In this section a further analysis of the workflow used to combine the augmentation techniques is presented, in order to provide a better insight of the dataset created and used in this work.

As mentioned above, the workflow was designed in that way so that the dataset created to train the network represents a generalized and realistic problem. By using the three combinations of the augmentation techniques, three different subsets are created as shown in Fig. 17, that form the main dataset used in this work. The first combination as described earlier uses only the augmentation techniques designed to simulate off-axis iris images. This process is used twice for each dataset creating the off-axis iris subset. The second combination uses only the augmentation techniques that simulate unconstrained conditions. This process is used once for each dataset consisting thus the unconstrained condition subset. Finally, the third combination, uses the augmentation techniques simulating the off-axis iris images and unconstrained conditions. This process is used twice for each dataset formulating the off-axis & unconstrained condition iris subset. Bath800 and CASIA Thousand combined consist of around 50.000 samples. With the use of the described workflow, 250.000 augmented samples are created. The off-axis iris subset is 100.000 samples, the unconstrained condition subset is 50.000 samples and 100.000 more samples from the unconstrained condition and off-axis iris subset. With the addition of the 50.000 original samples from Bath800 and CASIA Thousand, the final dataset used consists of 300.000 samples. In Table 1, a further analysis is presented describing the percentage of samples, with each augmentation technique or their combination, to the dataset.

Regarding the UBIRIS v2 as stated above, it consists of samples acquired in unconstrained conditions and therefore there is not a necessity of augmenting the samples with the augmentation techniques simulating unconstrained conditions. The samples of UBIRIS v2 are augmented only with the use of the augmentation techniques simulating off-axis iris images. This procedure is operated twice, creating the off-axis iris subset of UBIRIS v2 which along with the original samples are used in this work.

Finally, the element of randomness introduced in the augmentation techniques as well as at the way they are combined as explained in the workflow plays an important role to the augmentation process. One instance of that is that the direction of the shadowing, stretching or tilting is chosen randomly for each image. Also, the volume that an augmentation technique is applied to an image is chosen randomly between a range of values. Additionally, as illustrated in the workflow a sample is augmented with one or more augmentation techniques combined in different ways. These three approaches make it possible that a variety of conditions are introduced into the dataset leading to a generalized solution and each time producing unique samples with different characteristics and distributions. Examples of augmented samples with the use of all three different workflow combinations and their corresponding ground truth are given in Fig. 18.

3. Network design & training

In this section the design of the network is presented along with a detailed comparison of its complexity with other CNN methods designed for the iris segmentation task followed by the procedure of training and fine-tuning.

3.1. Network design

For the segmentation task, a fully convolutional network inspired by Bazrafkan et al. (2018) is used, consisting of 10 layers. The network starts with a 3×3 kernel mapping the input (1

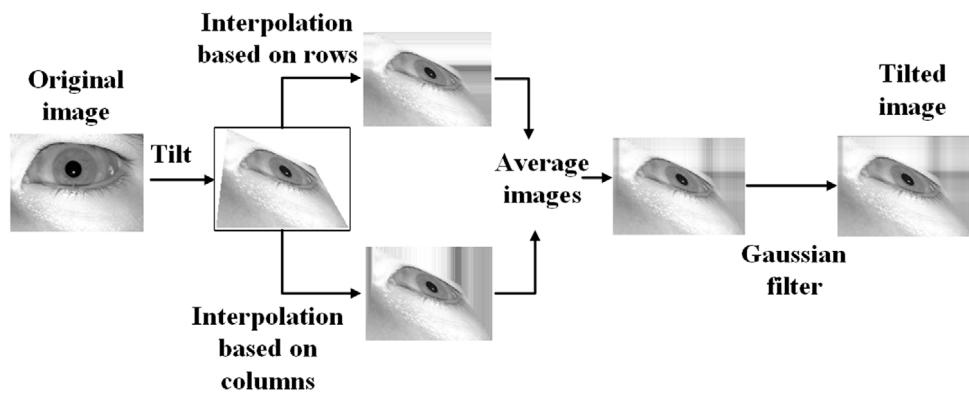


Fig. 15. Workflow of image tilting.

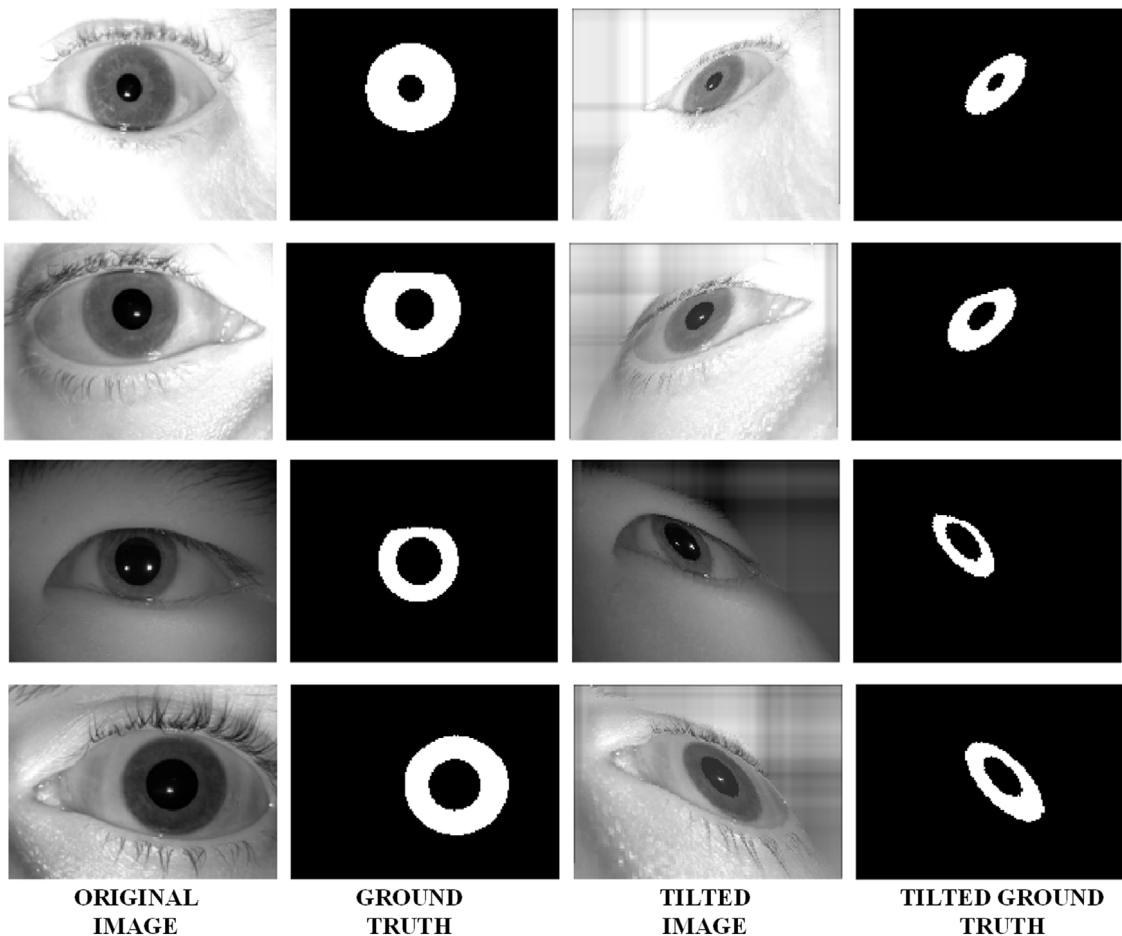


Fig. 16. Tilted samples and their corresponding segmentation map.

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 hidden convolutional layers, as well as, the number of channels and their activation function. Finally, at the output layer (1 channel), the kernel size is 3×3 , but in this layer, the sigmoid activation function is used. Pooling layers were not used as it was observed that the performance of the network's output was decreasing. The design of the network is illustrated in Fig. 19.

3.2. Complexity comparison of CNNs for iris segmentation

In this section, the complexity of several CNNs for iris segmentation will be compared with the proposed method. When referred to the complexity of a CNN, the main characteristics that one shall investigate is the number of parameters, the memory requirements for storing the parameters and the number of multiply-accumulate operations (MAC).

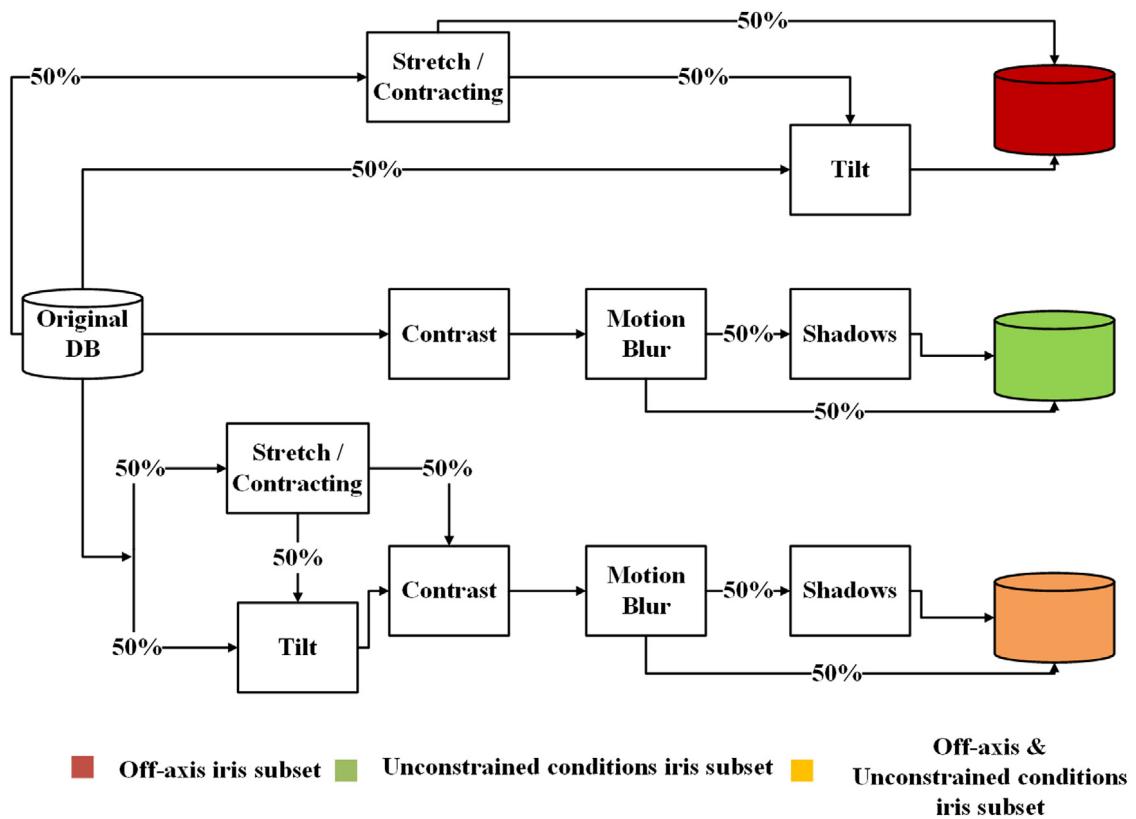


Fig. 17. Workflow of augmentation techniques.

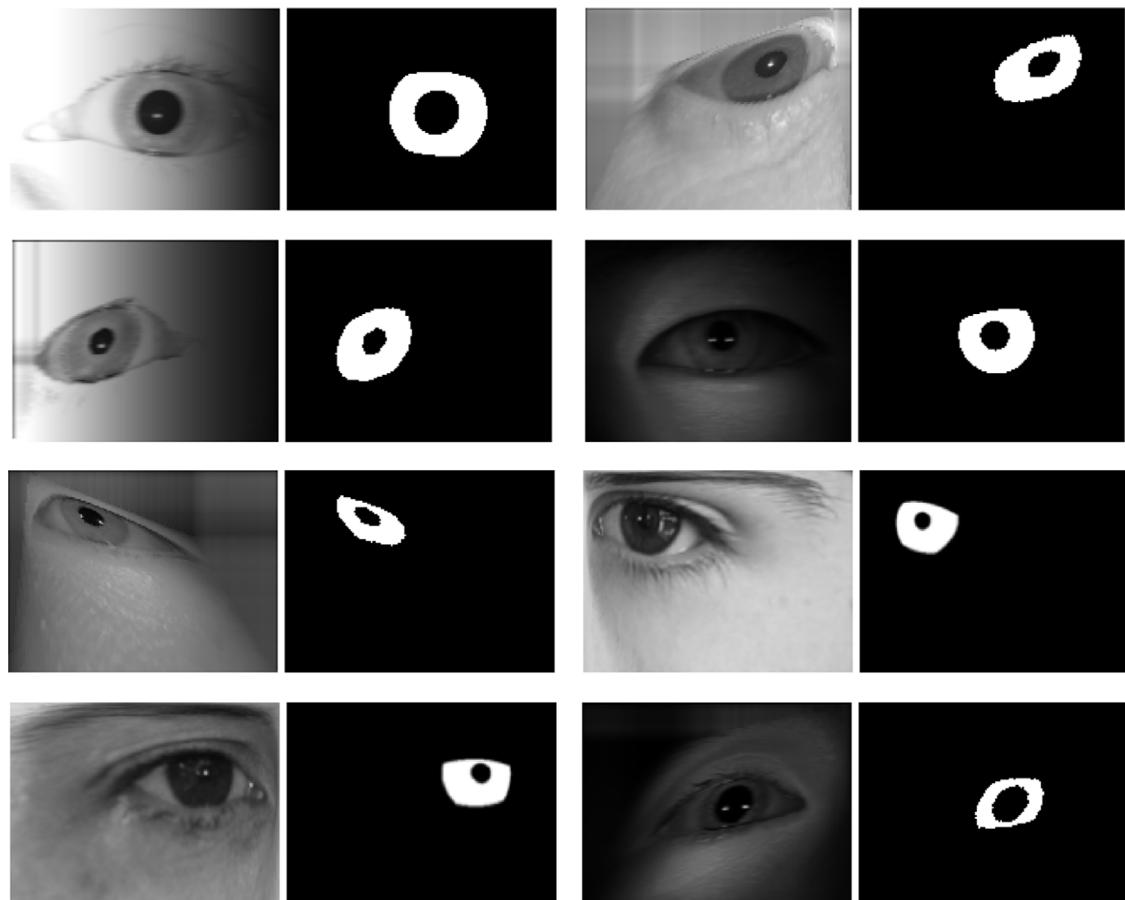
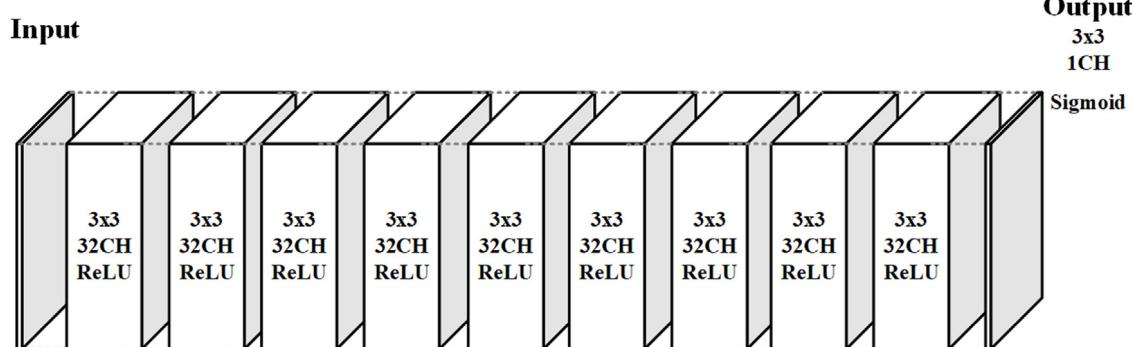


Fig. 18. Augmented samples and their corresponding ground truth.

Table 1

Percentage (%) of images with each augmentation technique or combination in the dataset. In this table the augmentation techniques are referred as Contrast reduction: Contrast, Motion blur: Blur, Shadows: Shadows, Spatial stretching/contracting: Warp and Tilting: Tilt.

Augmentation techniques	% of images	Dataset
Contrast & Blur	~8.5%	Unconstrained condition subset
Contrast & Blur & Shadows	~8.5%	
Warp	~8.5%	
Tilt	~16.5%	Off-axis subset
Warp & Tilt	~8.5%	
Warp & Contrast & Blur	~4%	Off-axis
Tilt & Contrast & Blur	~8.5%	&
Tilt & Contrast & Blur & Shadows	~8.5%	Unconstrained conditions subset
Warp & Contrast & Blur & Shadows	~4%	
Warp & Tilt & Contrast & Blur	~4%	
Warp & Tilt & Contrast & Blur & Shadows	~4%	
No augmentation	~16.5%	Original subset

**Fig. 19.** Network Design.

It is common practice that when an architecture of CNN is designed, that deep and large structures are favoured thus increasing the possibility of solving the investigated problem or promise higher performance from a smaller size CNN. Selecting a CNN with a deeper structure rather than a more compact structure, comes with some drawbacks such as increased training and execution time as well as generous memory requirements. There are cases, such as the proposed CNN, where a low complexity network can produce similar results as a high complexity network and as extension make it feasible to eliminate the downsides of a large CNN.

The proposed CNN consists of less than 75k parameters, requiring only 0.28 MB of memory to store the parameters and 1426.64M MAC for an input image with dimensions [120 × 160 × 1] [*width* × *height* × *channels*]. The SPDNN (Bazrafkan et al., 2018) consist of more than 1M parameters, requiring 35.26 MB to store them and 13536.22M MAC for a smaller input image of dimensions [98 × 128 × 1]. Another high performance deep learning method, MFCNs (Liu et al., 2016) is of high complexity and memory requirements, with 21M parameters, needing 82.56 MB of memory to store them. The input dimension and MAC in this structure are not specified as the input image dimension is not fixed and the number of MAC is related to the dimension of the input image. The complexity characteristics of the methods mentioned are shown in Table 2.

In this section is presented the low complexity proposed network, with reduced memory requirements resulting into a more efficient solution which is compatible for deployment in embedded applications such as AR/VR headsets. Furthermore, in the evaluation of the proposed network in Section 5, is demonstrated that the low complexity network proposed in this work can obtain high performance iris segmentation results in both off-axis and frontal samples. The proposed CNN is outperforming

Table 2
Complexity of CNNs for iris segmentation.

Metrics	Methods		
	Proposed method	SPDNN	MFCNs
Total no. parameters	74.593	1.101.851	21.643.596
Parameters size	0.28 MB	35.26 MB	82.56 MB
Input size dimensions (<i>width</i> × <i>height</i> × <i>channels</i>)	120 × 160 × 1	96 × 128 × 1	N/A
Total MAC	1426.64M	13 536.22M	N/A

other methods in segmenting off-axis iris images. Also, despite the fact that the network is designed for segmenting off-axis iris images, the results reported in segmenting frontal iris images are comparable to the-state-of-the-art SPDNN method of higher complexity and memory requirements.

3.3. Training and fine-tuning

3.3.1. Training

The network is trained on the original and augmented samples of Bath800 and CASIA Thousand. The dataset is divided 70% for the training set, 20% for validation set and 10% for the test set.

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, to optimize the loss function. The training is done on a desktop computer with Nvidia GTX 1080 GPU. The executable of the trained network is available at⁷.

3.3.2. Fine-tuning

In this section the process of fine-tuning the original network with the UBIRIS v2 dataset is described. Fine-tuning is a concept of transfer learning. Transfer learning is a machine learning technique, where knowledge gain during training in one type of problem is used to train in another related task or domain.

The proposed model was trained on the augmented and original samples from Bath800 and CASIA Thousand. UBIRIS v2 differs from the other datasets in the fact that it consists of visible iris image while Bath800 and CASIA Thousand are taken in NIR domain. Obtaining high-performance segmentation results in visible iris samples requires training a new model from the beginning or either fine-tune a pre-trained network on a dataset with visible samples. As UBIRIS v2 is a small dataset, training a new model is not possible, therefore fine-tuning the parameters of the pre-trained network is more functional. The network is trained on NIR iris samples and therefore it is expected that the network transfers the information and tune the parameters on the UBIRIS v2 samples, as the context of the task and the datasets are similar.

Regarding the specifics of fine-tuning, the network is fine-tuned on the augmented and original samples of UBIRIS v2. The dataset is divided 70% for the training set, 20% for validation set and 10% for the test set. 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 5e-5, beta1 and beta2 equal to 0.9 and 0.999 respectively, to optimize the loss function. The tuning is done on a desktop computer with Nvidia GTX 1080 GPU.

4. Results

The input of the network is a greyscale iris image of 1 channel with dimensions $[120 \times 160]$. The output of the network is a greyscale segmentation map with values between $[0, 1]$ and of the same size and channels as the input. The binary segmentation 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 for the Bath800, CASIA Thousand which are datasets containing NIR images. Regarding UBIRIS v2 which contains visible samples, after fine-tuning the network to the dataset, the threshold with value 0.4 is selected. The output of the proposed model for the different datasets is shown in Figs. 20–22.

5. Evaluation

Several metrics are used to evaluate the proposed method and conduct a detailed comparison with several segmentation methods of the literature. The metrics used in this work are: accuracy, sensitivity, specificity, precision, NPV and F1-score. More information about these metrics can be found in [Bazrafkan et al. \(2018\)](#). Two main experiments have been used to evaluate the performance of the proposed network:

- (1) Evaluate the proposed network on the off-axis augmented samples:

The network is trained on the original and augmented samples of Bath800 and CASIA Thousand. The network is tested on the off-axis augmented samples of Bath800, CASIA Thousand and UBIRIS v2. These are the off-axis subset and off-axis with unconstrained condition subset for Bath800 and CASIA Thousand and the off-axis subset for UBIRIS v2, as described in Section 2.3.3. In continuance it is compared with the segmentation results on these samples from the methods: SPDNN ([Bazrafkan et al., 2018](#)), IrisSeg

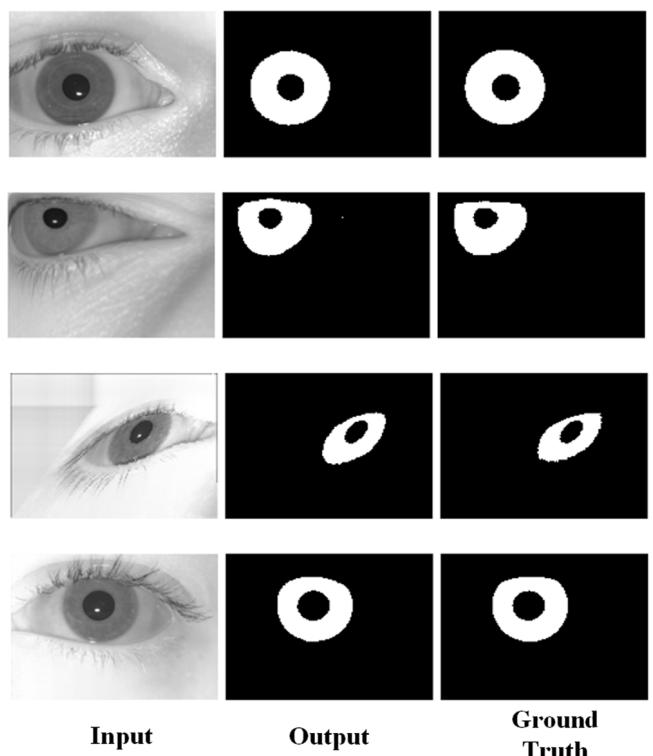


Fig. 20. Output of the network for the augmented off-axis and original samples of Bath800.

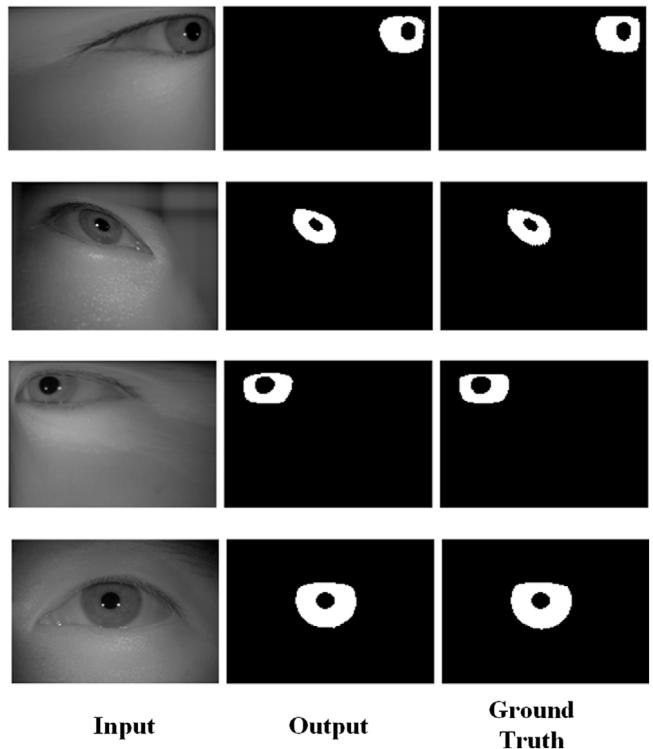


Fig. 21. Output of the network for the augmented off-axis and original samples of CASIA Thousand.

([Gangwar, Joshi, Singh, Alonso-Fernandez, & Bigun, 2016](#)) and OSIRIS ([Othman, Dorizzi, & Garcia-Salicetti, 2016](#)). The test set of the augmented samples is used to test the network and the other methods.

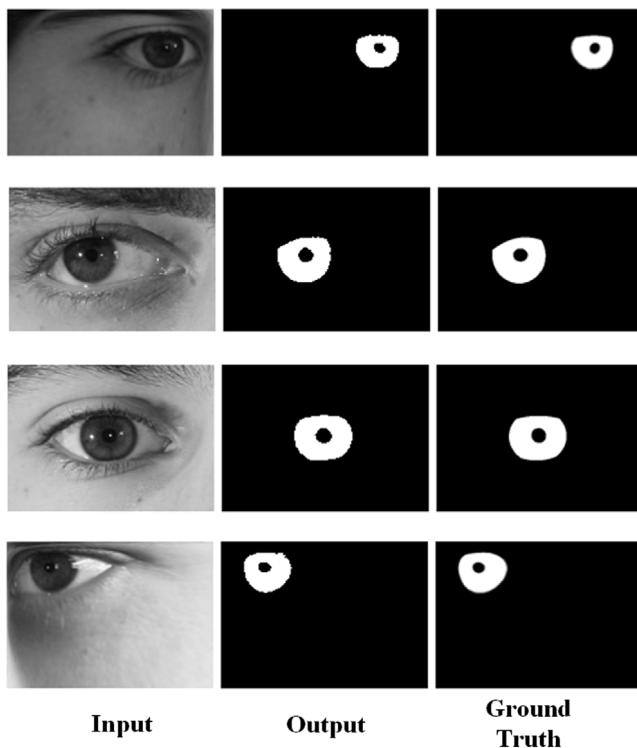


Fig. 22. Output of the network for the augmented off-axis and original samples of UBIRIS v2.

(2) Evaluate the network on the original samples from the datasets:

The network is tested on the original samples of Bath800, CASIA Thousand and UBIRIS v2, which consist of frontal iris samples. The test set of these datasets is used for testing the proposed method. The results of the proposed network are compared extensively with the state-of-the-art SPDNN on the Bath800, CASIA Thousand and UBIRIS v2. Furthermore, the results of the network are compared with the available results from other segmentation methods of the literature.

The results presented on UBIRIS v2 are the results of the original network after tuning.

5.1. Evaluation and comparison on off-axis augmented samples

In this section the proposed method is tested on the off-axis augmented samples. These samples are the combination of the off-axis subset and the off-axis with unconstrained condition subset for Bath800 and CASIA Thousand and the off-axis iris subset for UBIRIS v2. The test sets from the datasets are used for the testing stage.

5.1.1. Evaluation

The proposed network produces high performance results in the datasets Bath800 and CASIA Thousand. This is expected as the network is trained on them. On UBIRIS v2 the network is able to provide accurate segmentation results but is not able to perform at the same level as on Bath800 and CASIA Thousand. The samples of UBIRIS v2 are taken in visible spectrum and therefore the distribution differs. The proposed network with tuning is able to produce high segmentation results showing that the CNN is able to adopt to a similar task but with different distribution.

Table 3

Comparison of the proposed method with other segmentation methods on the off-axis augmented samples of Bath800. A higher value for μ and lower for σ is desired.

Metrics	Bath800				
	Proposed method	SPDNN	IrisSeg	OSIRIS	
Accuracy	μ	99.22%	97.03%	96.10%	95.86%
	σ	0.62%	1.96%	3.53%	2.80%
Sensitivity	μ	92.98%	58.71%	67.26%	62.16%
	σ	8.7%	38.04%	21.82%	35.72%
Specificity	μ	99.62%	99.15%	98.02%	98.00%
	σ	0.38%	0.86%	3.53%	2.37%
Precision	μ	93.97	80.34%	75.88%	67.68%
	σ	7.41%	19.32%	21.18%	24.11%
NPV	μ	99.52%	97.74%	97.79%	97.60%
	σ	0.57%	2.14%	1.72%	2.24%
F1-score	μ	93.21%	59.90%	68.63%	59.54%
	σ	7.70%	35.76%	19.51%	31.78%

In the datasets that the network is trained on, the accuracy and the F1-score and sensitivity measurements are higher showing high quality in returning true results and more consistent segmentations in comparison with UBIRIS v2. The same applies with the sensitivity and NPV metrics showing that the network is able to rule-out non-iris pixels more effectively in the trained datasets than the dataset that it was tuned on. Although, the network has higher performance in precision and specificity on the UBIRIS v2 dataset, showing greater capability in returning real iris pixels in the UBIRIS v2 dataset rather than the Bath800 and CASIA Thousand. The results are shown in Tables 3–5.

5.1.2. Comparison with SPDNN, IrisSeg and OSIRIS

The proposed method is designed for segmenting low quality off-axis iris images as acquired from an AR/VR device. The proposed method is compared with the SPDNN, IrisSeg and OSIRIS on the test set of the augmented off-axis samples. The selection of these algorithms is based on their availability. Furthermore, the SPDNN is a state-of-the-art segmentation method specialized on low quality iris images and IrisSeg and OSIRIS are well-established methods with high performance in the iris segmentation task. The SPDNN is trained on the original and augmented samples of Bath800 and CASIA Thousand and tuned on UBIRIS v2. The augmented samples used in their work are representing unconstrained scenarios. The SPDNN is a network with high capacity and large number of parameters as analysed earlier.

The SPDNN when tested on the off-axis augmented samples is able to provide overall good results in accuracy and specificity and average results in precision. The performance of the SPDNN is low in the sensitivity and F1-score measurements. The proposed network is outperforming the SPDNN in all the evaluation metrics showing higher results and ability to segment off-axis iris samples as appear when acquired from a user-facing camera on AR/VR device. In regard to IrisSeg and OSIRIS there are not able to provide high segmentation results for the augmented off-axis samples. The low performance results of IrisSeg and OSIRIS are due to the fact that the augmented samples that used are challenging as they simulate off-axis iris images in unconstrained conditions. In addition, IrisSeg and OSIRIS were not able to provide a segmentation in many cases. The results included for IrisSeg and OSIRIS are only for the images that the algorithms were able to provide a segmentation. The results are given in Tables 3–5.

Table 4

Comparison of the proposed method with other segmentation methods on the off-axis augmented samples of CASIA Thousand. A higher value for μ and lower for σ is desired.

Metrics	CASIA Thousand			
	Proposed method	SPDNN	IrisSeg	OSIRIS
Accuracy	μ 99.40%	97.75%	96.7%	95.81%
	σ 0.56%	1.66%	5.52%	2.49%
Sensitivity	μ 90.64%	49.36%	69.67%	36.34%
	σ 11.14%	43.15%	27.13%	38.40%
Specificity	μ 99.77%	99.43%	97.90%	98.60%
	σ 0.29%	0.9%	5.62%	2.13%
Precision	μ 94.17%	75.89%	74.37%	48.42%
	σ 7.87%	28.26%	27.97%	34.48%
NPV	μ 99.59%	98.27%	98.63%	97.07%
	σ 0.49%	1.64%	1.19%	2.14%
F1-score	μ 91.93%	49.40%	69.00%	35.83%
	σ 9.66%	41.44%	26.72%	34.31%

Table 5

Comparison of the proposed method with other segmentation methods on the off-axis augmented samples of UBIRIS v2. A higher value for μ and lower for σ is desired.

Metrics	UBIRIS v2			
	Proposed method	SPDNN	IrisSeg	OSIRIS
Accuracy	μ 98.83%	97.94%	87.17%	92.96%
	σ 1.16%	1.84%	9.10%	5.31%
Sensitivity	μ 83.89%	60.17%	27.06%	24.11%
	σ 10.48%	34.20%	23.51%	29.04%
Specificity	μ 99.77%	99.75%	91.03%	97.58%
	σ 0.46%	0.73%	9.30%	4.15%
Precision	μ 95.26%	93.78%	24.31%	39.11%
	σ 9.87%	15.51%	29.21%	38.34%
NPV	μ 98.94%	98.01%	94.97%	95.15%
	σ 1.12%	1.88%	4.22%	4.33%
F1-score	μ 88.72%	66.35%	21.58%	23.58%
	σ 10.62	35.49%	22.50%	29.66%

5.2. Evaluation and comparison on the frontal iris-region samples

In this section the proposed method is evaluated and compared on the frontal original samples of Bath800, CASIA Thousand and UBIRIS v2, which consist of frontal iris samples. It is worthwhile to note that the proposed technique is designed for segmenting off-axis consumer level iris images. Despite that, the experiments below are carried out in order to conduct a fair comparison with the other methods on frontal images. Meanwhile the proposed method is giving the best results on segmenting the augmented off-axis samples.

5.2.1. Evaluation on the frontal iris-region samples

The proposed network is now tested on the original samples from Bath800, CASIA Thousand and UBIRIS v2. For this procedure the test sets of the datasets are used.

Similar outcomes with the one's on the evaluation of the proposed method on the off-axis iris samples are found in the evaluation of the original samples. The proposed network has higher performance in the datasets that the network is trained on, Bath800 and CASIA Thousand. Lower performance is reported on UBIRIS v2. The network accomplishes high accuracy results in all datasets showing that has high quality in returning true results. Moreover, in all datasets it returns high values in specificity and precision, meaning that the model performs well returning iris pixels. The sensitivity metric on Bath800 and CASIA Thousand is

Table 6

Testing the proposed method on the original samples of several datasets.

Metrics	Proposed method		
	Bath800	CASIA Thousand	UBIRIS v2
Accuracy	μ 99.13%	99.50%	98.92%
	σ 0.52%	0.36%	0.67%
Sensitivity	μ 94.90%	94.67%	88.38%
	σ 6%	4.33%	9.29%
Specificity	μ 99.56%	99.86%	99.71%
	σ 0.47%	0.16%	0.39%
Precision	μ 95.67%	97.39%	96.33%
	σ 6.33%	2.83%	7.22%
NPV	μ 99.49%	99.63%	99.10%
	σ 0.45%	0.34%	0.60%
F1-score	μ 95.17%	95.94%	91.46%
	σ 5.43%	2.89%	9.63%

high, showing the ability of the model in ruling out non-iris pixels accurately while in UBIRIS v2 the same metric has average performance. The same applies to the F1-score measurement showing that the network produces more consistent segmentations, both in finding iris and non-iris pixels in the datasets Bath800 and CASIA Thousand compared to UBIRIS v2. In Table 6 the results of the proposed network on the test sets of the original samples from Bath800, CASIA Thousand and UBIRIS v2 are presented.

5.2.2. Comparison with the SPDNN

The SPDNN is a sophisticated network, with state-of-the-art results in the iris segmentation task. Now as mentioned earlier the SPDNN was trained on samples of Bath800 and CASIA Thousand and tuned on the UBIRIS v2, as is the proposed method. The SPDNN is of high complexity with 14 times more number of parameters when compared to the proposed network. This is an aspect that should be considered in the comparison between these segmentation methods. Also, as mentioned earlier the proposed network is designed for segmenting off-axis iris images as captured by a user-facing camera on AR/VR device. The numerical results of the SPDNN (Bazrafkan et al., 2018) performance are reported as presented in their work.

5.2.2.1. Comparing results on Bath800, CASIA Thousand and UBIRIS v2. The proposed method shows higher accuracy than the SPDNN in the Bath800 dataset which implies better quality in returning true results. The performance in specificity of the proposed method is also higher than the SPDNN. However in the precision metric the SPDNN is performing better. That shows that both can perform well in returning iris pixels, with not one method being better than the other. The same applies in the ability of the methods in ruling out non-iris pixels, as in NPV the proposed method is performing better than the SPDNN while the SPDNN shows higher results from the proposed method in the sensitivity metric. On the other hand, a small advantage of the SPDNN over the proposed method is in the F1-score showing a better efficiency. Overall in Bath800 dataset there is not a clear advantage of one method over the other as the performance in the metrics is divided with the differences between them either in favour or against them being marginal. The proposed network is performing comparable with the SPDNN in the Bath800 dataset.

In regard with the CASIA Thousand dataset, the SPDNN shows a small advantage over the proposed method. The proposed method performs better in the specificity and precision metrics showing higher quality in returning iris pixels than the SPDNN. In the rest of the evaluation metrics the SPDNN is performing better than the proposed method. Nonetheless, generally the differences in performance are marginal.

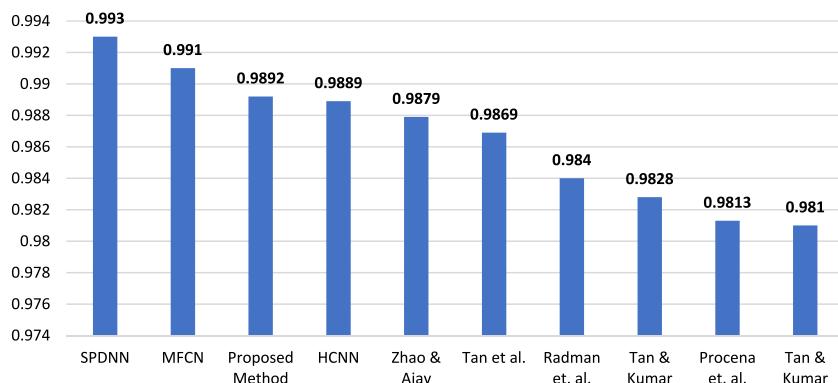


Fig. 23. Accuracy of the proposed method vs. other methods over the original UBIRIS v2 dataset.

On UBIRIS v2, the SPDNN performs better than the proposed method. The proposed method is performing better only in the specificity and precision metrics showing that is better on returning iris pixels than the SPDNN. In some metrics such as accuracy and NPV the difference is marginal showing that the proposed method is almost as good as the SPDNN in returning true results and in ruling out non-iris pixels. In the rest of the metrics there is a slight difference between the two methods, showing that the SPDNN is able to adopt better to the dataset that the methods are tuned utilizing thus the larger number of parameters of the SPDNN.

Overall, the proposed network and SPDNN performs similarly in the Bath800 and CASIA Thousand datasets, which are the datasets that were trained on. Therefore, showing that when trained the proposed network is comparable to the SPDNN despite that the complexity of the proposed network is at least an order of magnitude less than the SPDNN as analysed earlier. In the UBIRIS v2, where the proposed network and the SPDNN are tuned, the proposed method shows high results but the SPDNN still outperforms it, showing the ability to adopt better to a different dataset distribution, utilizing the higher complexity of its structure. The comparison between the two methods is shown in Table 7.

5.2.3. Comparison with state-of-the-art methods

In the following section the proposed method is compared to the most advance and state-of-the-art segmentation methods in the literature. First, accuracy over the challenging UBIRIS v2 dataset is compared with several methods. In continuance, it is evaluated and compared over CASIA Thousand and UBIRIS v2 in three important segmentation metrics: sensitivity, precision and F1-score with known segmentation methods.

5.2.3.1. Comparison of accuracy on UBIRIS v2. The accuracy of the proposed method is compared with state-of-the-art segmentation methods over UBIRIS v2. The state-of-the-art segmentation methods used in the comparison are the SPDNN (Bazrafkan et al., 2018), MFCN and HCNN from Liu et al. (2016) and Total Variation (TV) model utilized in Zhao and Ajay (2015); also an integrodifferential constellation followed by a curvature fitting model proposed in Tan, He, and Sun (2010), the HOG-SVM from Radman, Zainal, and Suandi (2017), and the random walker algorithm used in Tan and Kumar (2013). Moreover, the method proposed in Proenca (2010) where the sclera and iris regions are detected separately using neural networks as classifiers, and polynomial fitting is applied estimating the final iris region and finally the method from Tan and Kumar (2012) in which proposes a post-classification procedure including reflection and shadow removal and several refinements on pupil and eyelid localizations to get higher performance on iris segmentation task. The

accuracy of the proposed method compared with the aforementioned state-of-the-art methods on UBIRIS v2 dataset is presented in Fig. 23.

As illustrated, the proposed method has the third best performance compared with the state-of-the-art segmentation methods. The two methods that are performing better are: the SPDNN of Bazrafkan et al. (2018) and MFCN of Liu et al. (2016). However, these two methods are considerably more complex. As analysed in Section 3.2, SPDNN and MFCN contain 1M and 21M parameters respectively while the proposed network contains only 75k parameters.

Overall, the proposed method is the third best performing algorithm in the challenging dataset of UBIRIS v2 while its complexity is at least an order of magnitude less than the two methods that outperforms it, making the proposed method more suited for deployment in embedded applications.

5.2.3.2. Comparison on CASIA Thousand and UBIRIS v2. In this section the proposed method is compared with other known segmentation methods over three important metrics: sensitivity, precision and F1-score. Sensitivity measures the model's ability to rule out non-iris pixels, while precision measures the ability of the model to detect true iris pixels. F1-score is the harmonic average of these two metrics. The segmentation methods include CAHT, GST, IFFP, OSIRIS, and WAHET. The comparisons are made over the CASIA Thousand and UBIRIS v2 original datasets. The numerical results are initially presented at (Hofbauer et al., 2014). The metrics for each presented algorithm are calculated comparing the algorithms results with the ground truth. The comparisons are illustrated in Figs. 24–25. Comparing all methods on the high-quality CASIA Thousand dataset, the proposed method achieves the best performance on the F1-score and precision metrics, and second best for the sensitivity metric. Furthermore, on UBIRIS v2 where the samples are of low quality the proposed method gives higher results in all metrics compared to the other approaches. Although the proposed method is designed for segmenting off-axis iris samples, these results show that it performs well on frontal iris samples of high and low quality.

6. Conclusion

In this paper advanced data augmentation techniques are proposed to simulate off-axis iris samples as represented when captured by user-facing cameras on wearable AR/VR headsets, which enables us to propose a low-complexity neural network architecture, designed for deployment on embedded devices, targeting the segmentation of off-axis iris samples. The current network represents a proof of concept which will be integrated into hardware in future works. The quality of segmentation achieved by

Table 7

Comparison between the proposed method and the SPDNN on the original samples from several datasets. Green colour shows a better performance. Yellow shows a marginal difference in the performance and Orange a noteworthy difference in performance. A higher value for μ and lower for σ is desired.

Metrics		Bath800		CASIA Thousand		UBIRIS v2	
		Proposed Method	SPDNN	Proposed Method	SPDNN	Proposed Method	SPDNN
Accuracy	μ	99.13%	98.55%	99.50%	99.71%	98.92%	99.30%
	σ	0.52%	1.43%	0.36%	0.33%	0.67%	0.54%
Sensitivity	μ	94.90%	96.03%	94.67%	97.96%	88.38%	93.98%
	σ	6%	4.76%	4.33%	2.95%	9.29%	9.45%
Specificity	μ	99.56%	99.10%	99.86%	99.82%	99.71%	99.62%
	σ	0.47%	1.07%	0.16%	0.20%	0.39%	0.48%
Precision	μ	95.67%	96.05%	97.39%	97.13%	96.33%	94.88%
	σ	6.33%	4.46%	2.83%	3.10%	7.22%	5.40%
NPV	μ	99.49%	99.05%	99.63%	99.87%	99.10%	99.60%
	σ	0.45%	1.49%	0.34%	0.28%	0.60%	0.30%
F1-Score	μ	95.17%	95.93%	95.94%	97.50%	91.46%	93.90%
	σ	5.43%	3.88%	2.89%	2.51%	9.63%	9.70%

CASIA THOUSAND

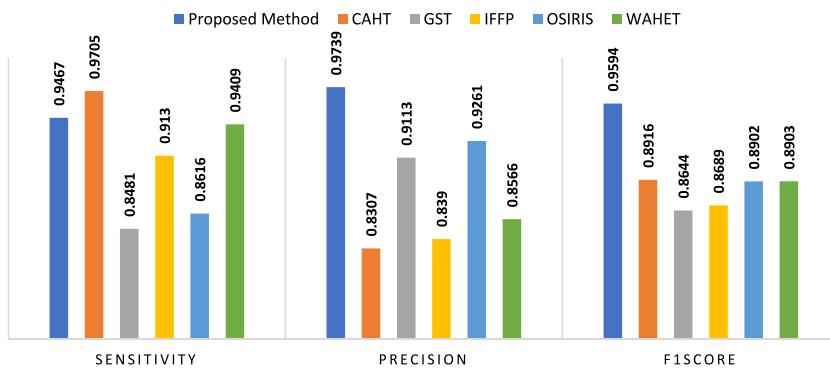


Fig. 24. Sensitivity, Precision, F1-score on the original samples of CASIA Thousand for the proposed method vs. five other methods.

UBIRIS V2

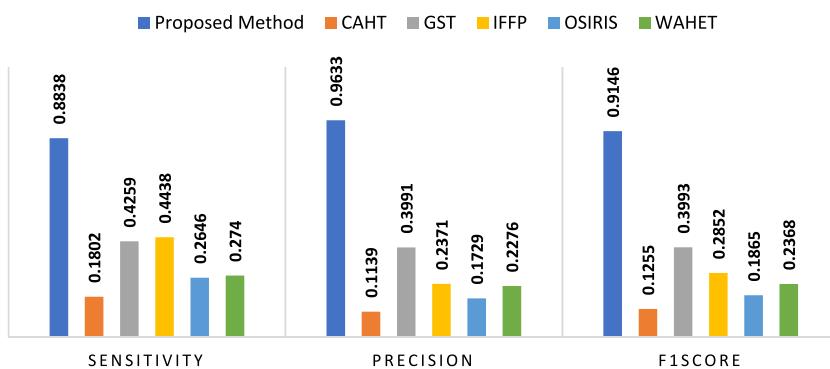


Fig. 25. Sensitivity, Precision, F1-score on the original samples of UBIRIS v2 for the proposed method vs. five other methods.

this network is evaluated and compared with state-of-the-art methods both for off-axis and frontal iris regions.

The proposed network's complexity is at least an order of magnitude less than other CNNs specifically designed for the iris segmentation task. Also, it has the best performance on segmenting the augmented off-axis iris samples. Further, the segmentation performance of this network on frontal iris samples from several public datasets, is comparable with the SPDNN network proposed by Bazrafkan et al. (2018) a state-of-the-art iris segmentation method. This performance is achieved even though the proposed network is of significantly smaller size and complexity and is trained for the task of segmenting off-axis iris samples. Due to its lightweight design and high performance in segmenting both off-axis and frontal iris samples and handling a range of input image qualities, the proposed network is well suited for general deployment on AR/VR devices.

Future work will focus on refinements in the network design and training/augmentation methodologies to improve performance on specific AR/VR headsets. As can be noted from the introduction, different devices will have user-facing cameras in a more limited set of locations and image acquisitions will be at varying NIR/wavelengths. In addition, the imaging pipeline on each camera module can have subtle effects on image quality.

Some practical examples of further research topics include developing an optimized CNN design based on SPDNN methods with a similar, or perhaps even smaller number of parameters that can achieve similar segmentation accuracy to our network. An additional further work includes the study of disease affected irises and the design of a CNN segmentation technique that is able to handle such iris images. Another future research direction is to build some device-specific datasets with iris images captured by the user-facing camera on several state-of-the-art AR/VR headsets. This will enable evaluation of the proposed segmentation method on practical off-axis iris samples. (At present it is not possible to gain low-level access to the imaging systems on the available devices to capture continuous image streams, but we have opened some discussions with device manufacturers and such access will hopefully be available in the near future as these devices continue to enter mainstream adoption.)

It is also expected to extend this work to apply these improved segmentation techniques to a number of full iris recognition pipelines to evaluate its effects on the reliability and robustness of near-view, off-axis iris recognition. The main challenge here is that the only off-axis recognition pipeline that we are aware of is proprietary. Again, we expect other algorithms will appear in the near future and hopefully some of these will be open-source or provide at least API-level access to system developers.

Acknowledgements

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).

References

- Abhyankar, A., Hornak, L., & Schuckers, S. (2005). Off-angle iris recognition using bi-orthogonal wavelet network system. In *Null* (pp. 239–244).
- Abhyankar, A., & Schuckers, S. (2006). Active shape models for effective iris segmentation. In *Biometric technology for human identification III*, Vol. 6202 (p. 62020H).
- Ackerman, E. (2013). Google gets in your face: Google glass offers a slightly augmented version of reality. *IEEE Spectrum*, 50(1), 26–29. <http://dx.doi.org/10.1109/MSPEC.2013.6395302>.
- Arsalan, M., Hong, H. G., Naqvi, R. A., Lee, M. B., Kim, M. C., Kim, D. S., et al. (2017). Deep learning-based iris segmentation for iris recognition in visible light environment. *Symmetry*, 9(11), 263.
- Arsalan, M., Naqvi, R. A., Kim, D. S., Nguyen, P. H., Owais, M., & Park, K. R. (2018). Irisdensenet: Robust iris segmentation using densely connected fully convolutional networks in the images by visible light and near-infrared light Camera sensors. *Sensors*, 18(5), 1501.
- Bakir, A., Chesler, G., & Torriente, M. de la (2016). Using touch ID for local authentication. *Internet of things with swift for iOS*.
- Bazrafkan, S., Thavalengal, S., & Corcoran, P. (2018). An end to end deep neural network for iris segmentation in unconstrained scenarios. *Neural Networks*, 106, 79–95. <http://dx.doi.org/10.1016/J.NEUNET.2018.06.011>.
- Bhorkar, G. (2017). A survey of augmented reality navigation. *Foundations and Trends® in Human-Computer Interaction*, 8(2–3), 73–272. <http://dx.doi.org/10.1561/1100000049>.
- Bowyer, K. W., Hollingsworth, K., & Flynn, P. J. (2008). Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding*, 110(2), 281–307.
- Bowyer, K. W., Hollingsworth, K. P., & Flynn, P. J. (2013). A survey of iris biometrics research: 2008–2010. In *Handbook of iris recognition*, Vol. 1 (pp. 15–54). Springer.
- Broussard, R. P., & Ives, R. W. (2009). Using artificial neural networks and feature saliency to identify iris measurements that contain the most discriminatory information for iris segmentation. In *Computational intelligence in biometrics: Theory, algorithms, and applications, 2009. CIB 2009. IEEE workshop on* (pp. 46–51).
- CASIA Iris Image Database (2019). (n.d.) Retrieved from <http://biometrics.idealtest.org/>.
- Cave, A. (2015). Why Google Glass Flopped. Retrieved from <http://www.forbes.com/sites/andrewcave/2015/01/20/a-failure-of-leadership-or-design-why-google-glass-flopped/#24c650d3556a>.
- Chan, P., Halevi, T., & Memon, N. (2015). *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics, Glass OTP: Secure and convenient user authentication on google glass*. http://dx.doi.org/10.1007/978-3-662-48051-9_22.
- Chauhan, J., Asghar, H. J., Kâafar, M. A., & Mahanti, A. (2016). Gesture-based continuous authentication for wearable devices: the google glass Case. In *14th international conference on applied cryptography and network security*.
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2014). Semantic image segmentation with deep convolutional nets and fully connected crfs. ArXiv Preprint [ArXiv:1412.7062](https://arxiv.org/abs/1412.7062).
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(4), 834–848.
- Chen, L.-C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. ArXiv Preprint [ArXiv:1706.05587](https://arxiv.org/abs/1706.05587).
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 801–818).
- Cherapau, I., Muslukhov, I., Arachchilage, N. A. G., & Beznosov, K. (2015). On the impact of touch ID on iphone passcodes. In *Symposium on usable privacy and security (SOUPS)* (pp. 257–276).
- Ching, K. W., & Singh, M. M. (2016). Wearable technology devices security and privacy vulnerability analysis. *International Journal of Network Security & Its Applications*, <http://dx.doi.org/10.5121/ijnsa.2016.8302>.
- Cognard, Timothée E., Goncharov, Alexander, Devaney, Nicholas, Dainty, Chris, & Corcoran, Peter (2018). undefined. (n.d.) A Review of Resolution Losses for AR/VR Foveated Imaging Applications. ieeexplore.ieee.org.
- Corcoran, P. M. (2013). Biometrics and consumer electronics: A brave new world or the road to dystopia?. *Consumer Electronics Magazine IEEE*, 2(2), 22–33.
- Corcoran, Peter (2016). The battle for privacy in your pocket [notes from the editor]. *IEEE Consumer Electronics Magazine*, 5(3), 3–36. <http://dx.doi.org/10.1109/MCE.2016.2558218>.
- Corcoran, P. M. (2017). A privacy framework for the internet of thing. In *2016 IEEE 3rd world forum on internet of things, WF-IoT 2016*. <http://dx.doi.org/10.1109/WF-IoT.2016.7845505>.
- Corcoran, Peter, Bigioi, P., & Thavalengal, S. (2015). Feasibility and design considerations for an iris acquisition system for smartphones. In *IEEE international conference on consumer electronics - Berlin, ICCE-Berlin (Vol. 2015-Febru* (pp. 164–167). IEEE, <http://dx.doi.org/10.1109/ICCE-Berlin.2014.7034328>.
- Corcoran, P., & Costache, C. (2016). Biometric technology and smartphones: A consideration of the practicalities of a broad adoption of biometrics and the likely impacts. *IEEE Consumer Electronics Magazine*, 5(2), 70–78. <http://dx.doi.org/10.1109/MCE.2016.2521937>.
- Cubuk, E. D., Zoph, B., Mane, D., Vasudevan, V., & Le, Q. V. (2018). Autoaugment: Learning augmentation policies from data. ArXiv Preprint [ArXiv:1805.09501](https://arxiv.org/abs/1805.09501).

- Darwaish, S. F., Moradian, E., Rahmani, T., & Knauer, M. (2014). Biometric identification on android smartphones. In *Procedia computer science*, Vol. 35 (pp. 832–841). <http://dx.doi.org/10.1016/j.procs.2014.08.250>.
- Daugman, J. (2007). New methods in iris recognition. *IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)*, 37(5), 1167–1175.
- Daugman, J. (2009). How iris recognition works. In *The essential guide to image processing* (pp. 715–739). Elsevier.
- De Luca, A., Hang, A., von Zezschwitz, E., & Hussmann, H. (2015). I feel like i'm taking selfies all day!. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems - CHI '15*. <http://dx.doi.org/10.1145/2702123.2702141>.
- Dorairaj, V., Schmid, N. A., & Fahmy, G. (2005). Performance evaluation of non-ideal iris based recognition system implementing global ICA encoding. In *Image processing, 2005. ICIP 2005. IEEE international conference on*, Vol. 3. III–285.
- Elise, B. (2014). Google Glass a Game-changing Application in the Realm of Cultural Tourism. Business Wire. Retrieved from http://search.proquest.com/docview/1635054301?accountid=14744%5Cnhttp://fama.us.es/search%sp;i?SEARCH=%5Cnhttp://pibserver.us.es/gtb/usuario_acceso.php?centro=USEG¢ro=%24USEG&d=1.
- Erbilek, M., Da Costa-Abreu, M. C., & Fairhurst, M. (2012). *Optimal configuration strategies for iris recognition processing*.
- Fox, B., & Felkey, B. (2013). Potential uses of google glass in the pharmacy. *Hospital Pharmacy*, 48(9), 783–784. <http://dx.doi.org/10.1310/hpj4809-783>.
- Gangwar, A., Joshi, A., Singh, A., Alonso-Fernandez, F., & Bigun, J. (2016). IrisSeg: A fast and robust iris segmentation framework for non-ideal iris images. In *2016 international conference on biometrics (ICB)* (pp. 1–8). IEEE.
- García-García, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., & García-Rodríguez, J. 2017. A review on deep learning techniques applied to semantic segmentation. ArXiv Preprint [ArXiv:1704.06857](https://arxiv.org/abs/1704.06857).
- Goode, A. (2014). Bring your own finger – how mobile is bringing biometrics to consumers. *Biometric Technology Today*, 2014(5), 5–9. [http://dx.doi.org/10.1016/S0969-4765\(14\)70088-8](http://dx.doi.org/10.1016/S0969-4765(14)70088-8).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Hammal, Z., Massot, C., Bedoya, G., & Caplier, A. (2005). In S. Singh, M. Singh, C. Apte, & P. Perner (Eds.), *Eyes segmentation applied to gaze direction and vigilance estimation BT - pattern recognition and image analysis* (pp. 236–246). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Hayes, A. (2016). My journey into glass: Talking about google glass with stakeholders in the glass explorer program. *IEEE Consumer Electronics Magazine*, 5(1), 102–105. <http://dx.doi.org/10.1109/MCE.2015.2484819>.
- He, X., & Shi, P. (2006). A novel iris segmentation method for hand-held capture device. In *International conference on biometrics* (pp. 479–485). Springer.
- Hofbauer, Heinz, Alonso-Fernandez, F., Bigun, J., & Uhl, A. (2016). Experimental analysis regarding the influence of iris segmentation on the recognition rate. *IET Biometrics*, 5(3), 200–211.
- Hofbauer, H., Alonso-Fernandez, F., Wild, P., Bigun, J., & Uhl, A. (2014). A ground truth for iris segmentation. In *2014 22nd international conference on pattern recognition* (pp. 527–532). <http://dx.doi.org/10.1109/ICPR.2014.101>.
- Huang, Y.-P., Luo, S.-W., & Chen, E.-Y. (2002). An efficient iris recognition system. In *Machine learning and cybernetics, 2002. Proceedings. 2002 international conference on*, Vol. 1 (pp. 450–454). IEEE.
- Jalilian, E., Uhl, A., & Kwitt, R. (2017). Domain adaptation for CNN based Iris segmentation. *BIOSIG 2017*.
- Jan, F. (2017). Segmentation and localization schemes for non-ideal iris biometric systems. *Signal Processing*, 133, 192–212.
- Jiang, Z., Yuan, Y., & Wang, Q. (2018). Contour-aware network for semantic segmentation via adaptive depth. *Neurocomputing*, 284, 27–35.
- Jillela, R., & Ross, A. A. (2013). Methods for iris segmentation. In *Handbook of Iris recognition* (pp. 239–279). Springer.
- Khan, T. M., Khan, M. A., Malik, S. A., Khan, S. A., Bashir, T., & Dar, A. H. (2011). Automatic localization of pupil using eccentricity and iris using gradient based method. *Optics and Lasers in Engineering*, 49(2), 177–187.
- Koh, J., Govindaraju, V., & Chaudhary, V. (2010). A robust iris localization method using an active contour model and hough transform. In *Pattern recognition (ICPR), 2010 20th international conference on* (pp. 2852–2856).
- Kress, B., Saeedi, E., & Brac-de-la Perriere, V. (2014). 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*, Vol. 9202 (p. 92020D). <http://dx.doi.org/10.1117/12.2064351>.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097–1105).
- Lakra, A., Tripathi, P., Keshari, R., Vatsa, M., & Singh, R. (2018). Segdensenet: Iris segmentation for pre-and-post Cataract surgery. In *2018 24th international conference on pattern recognition (ICPR)* (pp. 3150–3155). IEEE.
- Lateef, F., & Ruichek, Y. (2019). Survey on semantic segmentation using deep learning techniques. *Neurocomputing*.
- Lemley, J., Bazrafkan, S., & Corcoran, P. (2017). Smart augmentation learning an optimal data augmentation strategy. *IEEE Access*, 5, 5858–5869. <http://dx.doi.org/10.1109/ACCESS.2017.2696121>.
- Li, X. (2006). Modeling intra-class variation for nonideal iris recognition. In *International conference on biometrics* (pp. 419–427).
- Lili, P., & Mei, X. (2005). The algorithm of iris image preprocessing. In *Automatic identification advanced technologies, 2005. Fourth IEEE workshop on* (pp. 134–138). IEEE.
- Linao, M. (2016). *The present and future of VR/AR: Applications in education, gaming, commerce, and industry*.
- Liu, N., Li, H., Zhang, M., Liu, J., Sun, Z., & Tan, T. (2016). Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In *Biometrics (ICB), 2016 International conference on* (pp. 1–8).
- Liu, Y., Yuan, S., Zhu, X., & Cui, Q. (2003). A practical iris acquisition system and a fast edges locating algorithm in iris recognition. In *IEEE instrumentation and measurement technology conference proceedings*, Vol. 1 (pp. 166–169).
- Mann, S. (2001). Fundamental issues in mediated reality, wearcomp, and camera-based augmented reality. In *Fundamentals of wearable computers and augmented reality* (pp. 295–328). Lawrence Erlbaum Associates, Inc., [http://dx.doi.org/10.1061/\(ASCE\)0733-947X\(2005\)131:3\(169\)](http://dx.doi.org/10.1061/(ASCE)0733-947X(2005)131:3(169)).
- Mann, S. (2004). Continuous lifelong capture of personal experience with EyeTap. In *Proceedings of the 1st ACM workshop on continuous archival and retrieval of personal experiences - CARPE'04* (pp. 1–21). <http://dx.doi.org/10.1145/1026653.1026654>.
- Mann, S. (2013). Steve mann: My augmented life. *IEEE Spectrum*, 1–6.
- Mann, S., & Fung, J. (2002). Eyetap devices for augmented, deliberately diminished, or otherwise altered visual perception of rigid planar patches of real-world scenes. *Presence: Teleoperators & Virtual Environments*, 11(2), 158–175. <http://dx.doi.org/10.1162/1054746021470603>.
- MIRLIN (2019). (n.d.) Retrieved from <https://www.fotonation.com/products/biometrics/iris-recognition/>.
- Muensterer, O. J., Lacher, M., Zoeller, C., Bronstein, M., & Kübler, J. (2014). Google glass in pediatric surgery: An exploratory study. *International Journal of Surgery*, 12(4), 281–289. <http://dx.doi.org/10.1016/j.ijsu.2014.02.003>.
- Othman, N., Dorizzi, B., & Garcia-Salicetti, S. (2016). OSIRIS: An open source iris recognition software. *Pattern Recognition Letters*, 82, 124–131.
- Peng, G., Zhou, G., Nguyen, D. T., Qi, X., Yang, Q., & Wang, S. (2017). Continuous authentication with touch behavioral biometrics and voice on wearable glasses. *IEEE Transactions on Human-Machine Systems*, <http://dx.doi.org/10.1109/THMS.2016.2623562>.
- Perez, L., & Wang, J. (2017). The effectiveness of data augmentation in image classification using deep learning. ArXiv Preprint [ArXiv:1712.04621](https://arxiv.org/abs/1712.04621).
- Prabhakar, S., Pankanti, S., & Jain, A. K. (2003). Biometric recognition: security and privacy concerns. *IEEE Security & Privacy*, 1(2), 33–42. <http://dx.doi.org/10.1109/MSECP.2003.1193209>.
- Proenca, H. (2010). Iris recognition: On the segmentation of degraded images acquired in the visible wavelength. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8), 1502–1516.
- Proenca, H., Filipe, S., Santos, R., Oliveira, J., & Alexandre, L. A. (2010). The ubiris v2: A database of visible wavelength iris images captured on-the-move and at-a-distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8), 1529–1535.
- Proença, H., & Alexandre, L. A. (2010). Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. *Image and Vision Computing*, 28(1), 202–206.
- Quinn, G. W., Grother, P. J., Ngan, M. L., & Matey, J. R. (2013). IREX IV: part 1, evaluation of iris identification algorithms.
- Radman, A., Zainal, N., & Suandi, S. A. (2017). Automated segmentation of iris images acquired in an unconstrained environment using HOG-svm and growcut. *Digital Signal Processing*, 64, 60–70.
- Rakshit, S. (2007). *Novel methods for accurate human Iris recognition*. University of Bath.
- Ring, T. (2015). Spoofing: are the hackers beating biometrics?. *Biometric Technology Today*, 2015(7), 5–9. [http://dx.doi.org/10.1016/S0969-4765\(15\)30119-3](http://dx.doi.org/10.1016/S0969-4765(15)30119-3).
- Rompapas, D. C., Rovira, A., Ikeda, S., Plopski, A., Taketomi, T., Sandor, C., et al. (2017). Eyear: Refocusable augmented reality content through eye measurements. In *Adjunct proceedings of the 2016 IEEE international symposium on mixed and augmented reality, ISMAR-Adjunct 2016*. <http://dx.doi.org/10.1109/ISMAR-Adjunct.2016.0108>.
- Russakovs, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211–252.
- Salamon, J., & Bello, J. P. (2017). Deep convolutional neural networks and data augmentation for environmental sound classification. *IEEE Signal Processing Letters*, 24(3), 279–283.
- Samangouei, P., Patel, V. M., & Chellappa, R. (2017). Facial attributes for active authentication on mobile devices. *Image and Vision Computing*, <http://dx.doi.org/10.1016/j.imavis.2016.05.004>.
- Schlüter, J., & Grill, T. (2015). Exploring data augmentation for improved singing voice detection with neural networks. In *ISMIR* (pp. 121–126).

- Schreinemacher, M. H., Graafland, M., & Schijven, M. P. (2014). Google glass in surgery. *Surgical Innovation*, <http://dx.doi.org/10.1177/1553350614546006>.
- Shah, S., & Ross, A. (2009). Iris segmentation using geodesic active contours. *IEEE Transactions on Information Forensics and Security*, 4(4), 824–836.
- Shejin, Thavalengal, & Corcoran, P. (2016). User authentication on smartphones: Focusing on iris biometrics. *IEEE Consumer Electronics Magazine*, 5(2), 87–93. <http://dx.doi.org/10.1109/MCE.2016.2522018>.
- Shijie, J., Ping, W., Peiyi, J., & Siping, H. (2017). Research on data augmentation for image classification based on convolution neural networks. In *2017 Chinese automation congress (CAC)* (pp. 4165–4170). IEEE.
- Starner, T., Mann, S., Rhodes, B., Levine, J., Healey, J., Kirsch, D., et al. (1997). Augmented reality through wearable computing. *Presence: Teleoperators and Virtual Environments*, 6(4), 386–398. <http://dx.doi.org/10.1162/pres.1997.6.4.386>.
- Tan, T., He, Z., & Sun, Z. (2010). Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. *Image and Vision Computing*, 28(2), 223–230.
- Tan, C.-W., & Kumar, A. (2012). Unified framework for automated iris segmentation using distantly acquired face images. *IEEE Transactions on Image Processing*, 21(9), 4068–4079.
- Tan, C.-W., & Kumar, A. (2013). Towards online iris and periocular recognition under relaxed imaging constraints. *IEEE Transactions on Image Processing*, 22(10), 3751–3765.
- Tang, F., Aimone, C., Fung, J., Marjan, A., & Mann, S. (2002). 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* (pp. 267–268). <http://dx.doi.org/10.1109/ISMAR.2002.1115106>.
- Taylor, L., & Nitschke, G. (2017). Improving deep learning using generic data augmentation. ArXiv Preprint [ArXiv:1708.06020](https://arxiv.org/abs/1708.06020).
- Thavalengal, Shejin, Andorko, I., Drimbarean, A., Bigioi, P., & Corcoran, P. (2015). Proof-of-concept and evaluation of a dual function visible/NIR camera for iris authentication in smartphones. *IEEE Transactions on Consumer Electronics*, 61(2), 137–143. <http://dx.doi.org/10.1109/TCE.2015.7150566>.
- Thavalengal, S., Bigioi, P., & Corcoran, P. (2015a). Evaluation of combined visible/NIR camera for iris authentication on smartphones. In *2015 IEEE conference on computer vision and pattern recognition workshops (CVPRW)* (pp. 42–49). <http://dx.doi.org/10.1109/CVPRW.2015.7301318>.
- Thavalengal, S., Bigioi, P., & Corcoran, P. (2015b). Iris authentication in handheld devices - considerations for constraint-free acquisition. *IEEE Transactions on Consumer Electronics*, 61(2), 245–253. <http://dx.doi.org/10.1109/TCE.2015.7150600>.
- Thavalengal, Shejin, Bigioi, P., & Corcoran, P. (2016). Efficient segmentation for multi-frame iris acquisition on smartphones. In *2016 IEEE international conference on consumer electronics (ICCE) (2016 ICCE)* (pp. 202–203). Las Vegas, USA.
- Timekeeper (2017). The promise of augmented reality. *The Economist*.
- Tipton, Stephen J., White II, Daniel J., Sershon, Christopher, & Choi, Young B. (2014). Ios security and privacy: Authentication methods, permissions, and potential pitfalls with touch id. *International Journal of Computer and Information Technology*, 3(3), 482–489.
- Varkarakis, V., Bazrafkan, S., & Corcoran, P. (2018). A deep learning approach to segmentation of distorted iris regions in head-mounted displays. In *2018 IEEE games, entertainment, media conference (GEM)* (pp. 1–9). IEEE.
- Vazquez-Fernandez, E., & Gonzalez-Jimenez, D. (2016). Face recognition for authentication on mobile devices. *Image and Vision Computing*, <http://dx.doi.org/10.1016/j.imavis.2016.03.018>.
- Wang, Q., Gao, J., & Yuan, Y. (2018). Embedding structured contour and location prior in siamesed fully convolutional networks for road detection. *IEEE Transactions on Intelligent Transportation Systems*, 19(1), 230–241.
- WaveLab (2019). (n.d.) No Title..
- Wildes, R. P. (1997). Iris recognition: an emerging biometric technology. *Proceedings of the IEEE*, 85(9), 1348–1363.
- Yadav, D. K., Ionascu, B., Ongole, S. V. K., Roy, A., & Memon, N. (2015). *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics, Design and analysis of shoulder surfing resistant pin based authentication mechanisms on google glass*. http://dx.doi.org/10.1007/978-3-662-48051-9_21.
- Zhao, Z., & Ajay, K. (2015). An accurate iris segmentation framework under relaxed imaging constraints using total variation model. In *Proceedings of the IEEE international conference on computer vision* (pp. 3828–3836).