A Deep Learning Based Approach to Iris Sensor Identification

Ananya Zabin
LCSEE
West Virginia University
Morgantown, WV 26505
Email: az0019@mix.wvu.edu

Thirimachos Bourlai
MILAB, School of Electrical & Computer Engineering
University of Georgia
Athens, GA 30602
Email: Thirimachos.Bourlai@uga.edu

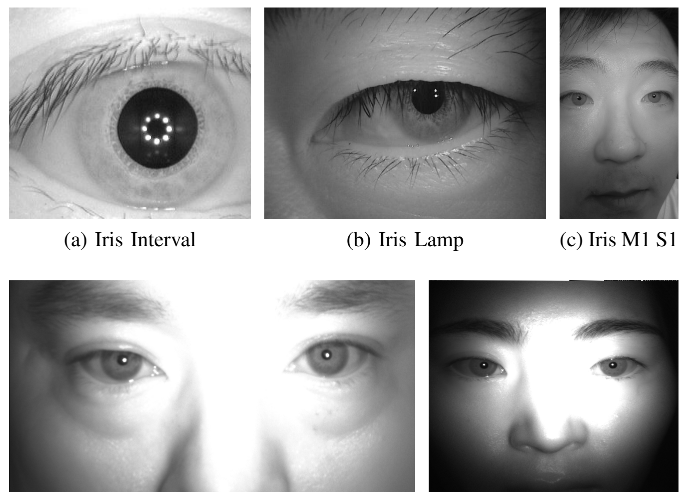
***Abstract*—An efficient iris sensor identification algorithm can be used in certain forensic applications, i.e. detecting mislabeled iris data at large scale iris datasets, and verifying the validity of the data origin of collected iris datasets that are available to be shared. Such knowledge can potentially increase the overall iris recognition system accuracy by offering the operator the option to match same-sensor or cross-sensor iris images. In either case the knowledge of the origin of the sensor used to collect these data, when not available, or the correction of mislabeled data, is expected to result in higher iris matching accuracy. Another benefit of iris sensor identification is that it can assist in improving the detection of fake iris data, i.e. when knowing the iris sensor, we can apply more appropriate models for fake detection that are tuned for a specific iris sensor. In this paper we propose an efficient deep learning-based iris recognition algorithm that is sensor inter-operable. Our approach utilizes a moderate amount of data and is adaptable to learning rate variations as well as variations of the amount of data used for training per class. Our proposed approach uses a set of iris datasets that include iris images captured at different standoff distances. We are using the original captured, dual eye, or periocular images rather than the iris itself, after detecting, segmenting, and normalizing the iris. Thus, the algorithm is efficient, fast, and less depended on additional algorithmic processes that can add computational complexity. Our proposed process includes transfer learning using iris images of higher quality via the utilization of a set of image quality metrics and achieves close to a hundred percent accuracy after cross-validation.**

***Index Terms*—Iris Sensor Identification, CNN, Transfer Learning, AlexNet, GoogLeNet, SGDM, Iris Sensor Inter-operability**

I. Introduction

Iris is among the most interesting biometric modalities. Iris recognition systems can be highly accurate as the iris features are unique from person to person. No two human iris patterns are same, not even in twins [1]. In many security related applications, where a highly efficient access control solution is needed, iris recognition systems are preferred to be used either independently or as part of a multi-modal biometric verification system.

While large number of iris recognition solutions are available, one of their challenges is that iris images can be captured using different imaging sensors. Thus, the issue of sensor interoperability can emerge, where some users may have enrolled using one iris sensor, while matching is performed with their live iris counterpart captured by another sensor. In such a case, we are performing iris cross-sensor matching, which can lower the efficiency of the iris recognition system in terms of its accuracy. This cannot be avoided if we do not know for sure the correct origin and thus, label of the available iris images we have available before matching. Thus, a capability to automatically determine the identity of an iris sensor from labelled or unlabeled iris images can be beneficial in many ways.



(a) Iris Interval (b) Iris Lamp (c) Iris M1 S1

(d) Iris M1 S2 (e) Iris M1 S3

Fig. 1: Large scale iris datasets when captured with multiple sensors - can suffer from cases of iris sensor mislabeling. In this example we show how onerous it is to manually check and verify the validity of the data origin of collected iris datasets when capturing using multiple iris sensors.

One of the challenges is that iris images are captured in the visible and near infrared bands via the same or different band-specific sensors. Thus, they differ due to the wavelength used, hardware related features and the illumination used while capturing the images [1], [2]. However, it is not always certain that we will have the correct label of each iris image captured. Another issue is that sensor features may differ even within

IEEE/ACM ASONAM 2020, December 7-10, 2020
978-1-7281-1056-1/20/$31.00 © 2020 IEEE

different camera models produced by same manufacturer [3]. So, in certain scenarios the amount of intra-class variation can be unavoidable, raising the requirement for sensor inter-operability.

Camera sensor recognition has also been studied for other biometric modalities such as fingerprint [4], voice, face [5] and one common conclusion is matched that the knowledge of the iris sensor used to capture an image helps the operator utilizing the proper models and algorithm [6]–[8].

In this paper we propose an efficient, deep learning-based, iris recognition algorithm that is sensor inter-operable. Our approach utilizes a moderate amount of data and is adaptable to learning rate variations as well as variations of the amount of iris images used for training per class. Our proposed approach uses a diverse iris dataset that include iris images captured at different standoff distances. Specifically, we are using a multi camera iris database collected from CASIA [9], and train the network sequentially with 300, 600, 900, 1200 and 1400 images per class. Also, we are using the original captured, dual eye, or periocular images rather than the iris itself, after detecting, segmenting, and normalizing the iris. The proposed algorithm is efficient, fast, and less depended on additional algorithmic processing that can add computational complexity. Also, our proposed process includes transfer learning - after we explored using different pre-trained models, we concluded that AlexNet and GoogLeNet result in the highest iris sensor identification accuracy. Finally, the usage of higher quality iris images via the utilization of a set of image quality metrics, yields close to a hundred percent accuracy after cross-validation.

The rest of the paper is organized as follows: Section 2 discusses previous works related to sensor interoperability, Section 3 and 4 discuss the methodological approach and experimental results, respectively. Limitations of this research is discussed in Section 5. Section 6 encompass the conclusions.

II. Related Work

Sensor interoperability research has been explored extensively for its vivid importance in forensics, bio-metrics verification, and others. While multiple solutions have been proposed, there have been several approaches adopted to find the exact camera model used starting from finding specific marks in the images [10]–[13].

Kirchner et al. [14] adopted a blind approach by extracting expressive texture-based features from the images that are used for training a classifier. The aim was to take advantage of the micro-pattern differences between the images captured with different camera sensors. In their algorithm they applied a high-pass filter to the data, which helped exploit the sensor related content that is useful for the texture classification approach.

Galdi et al. [13] used the sensor noise pattern for sensor identification. Such patterns are widely used in Photo response non-uniformity (PRNU) sensor recognition [15], [16]. Although this approach seems to be promising, it may be challenging to verify its efficacy as it is time consuming and needs a large number of images captured by individual sensors. This complicates the design of such an approach when we would also need to consider many sensors in combination with a large number of images needed from each sensor so that the sensor identification approach can work efficiently.

Celiktutan et al. [17] used different color bands to extract image quality metric (IQM) based features and combined them with Local Binary Patterns (LBP), extracted from the least-significant bit planes [18]. Gloe et al. [19] used additional color features and was able to improve the overall performance of their proposed system when compared to Xu et al. [20].

Recently, CNN based approaches have also been proposed to solve the sensor identification problem. Examples include the work at [21] with which Qian et al. [22] added a high-pass filter layer to extract features from the residuals of the images and, thus, reduce the overall complexity of the deep learning network.

III. Methodology

In this section, we discuss all the steps of our proposed approach as illustrated in Fig 3. First, each image is pre-processed by resizing it according to the input specification of the proposed network. After resizing, an image quality metrics is generated from all the images to choose the images needed to perform the experiment, as we choose the same number of images from each class for every experiment performed. Then, the images are divided into two sets: 90% for training and 10% for testing. Next, feature extraction is performed before training and validating the classifiers used to determine which sensor is used to capture the iris image in question.

A. Data Preprocessing

For this research, we are using five different classes of iris image datasets captured with five different cameras. Individual images in each class differ in size and number. For using a specific CNN architecture, all the images used in a dataset must be of same size, which is the exact input size of the network.

Thus, before the training phase starts, we need to make sure that all the images are converted to the exact input size of the network. For AlexNet the input images must be of size 227×227×3. Thus, we resize the data to be of the exact same input size to fulfil the requirement of the network selected.

B. Image Quality Metrics

We have been working with five sets of images from CASIA Iris datasets. Each set of images comes with different numbers of images, ranging from 1,400 to 20,000 images. For maintaining the regularity of the classification process, we choose same number of images per class. We run the experiment by choosing the images randomly as well as choosing better quality images to check whether the accuracy differs when compared to lower quality ones. Image quality can depend on several features of an image like noise, ringing, blurring and compression artifacts, thus being able to objectively measure the quality of an image is very important. As we choose

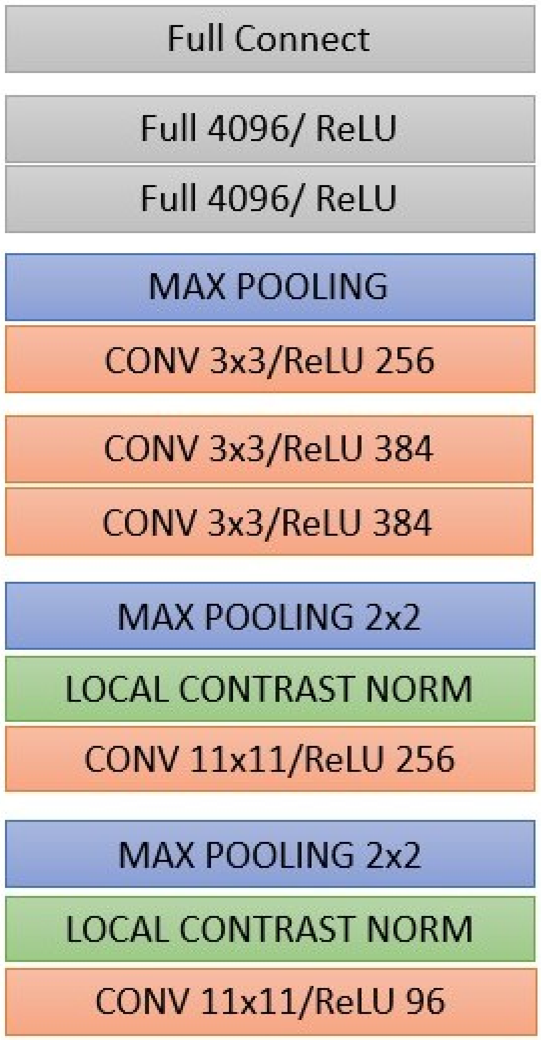


Fig. 2: Deep learning-based architecture followed to perform the experiment, where the first convolution layer consists 256 3×3 filters and the following two convolutions has 384 filters each. Max pooling 2×2 reduces the image size by one forth consisting only the most prominent features. The later convolution layers have 256 and 96 filters respectively where the size of the filters changes to 11×11.

to select the same number of images for each class, we measure the quality of images and choose the highest quality images per class to conduct the experiments. We generate no-reference-based image quality metrics with Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE). A brisque model is trained with a set of images with known distortion. We choose the highest quality images by generating a model, first, and then check the classification accuracy both with and without using a model. BRISQUE is limited to evaluating the quality of images with the same type of distortion. In BRISQUE there is a subjective quality score that accompanies the training images, which is predicted by using a support vector regression (SVR) model trained on the set of images with corresponding differential mean opinion score (DMOS) values. For generating these values, the compression artifacts, blurring, and noise are generated for the entire image database. The lower the BRISQUE score, the better the quality of that image is.

C. AlexNet Architecture

AlexNet was first introduced with the idea of implementing deep learning onto massive real-life datasets. Convolutional Neural Networks (CNN) being easy to control and train, has been the go-to model for object detection [23]. For vast datasets like ImageNet, it was needed a system to work smoothly with GPUs, provide with optimized training time and better performance.

It is eight layers deep (Fig: 2) convolutional neural network including input and output layers. Among those eight layers, five are convolutional layers and three fully-connected layers. There are some features which make AlexNet special and are new approaches to convolutional neural networks. Use of Rectified Linear Units (ReLU) instead of the tanh function reduces the training time by six times for the CIFAR-10 dataset. AlexNet was the first one introduced to train larger models by allowing for multi-GPU training.

To do so, it puts half of the model's neurons on one GPU and the other half on another. By introducing pooling overlap AlexNet could reduce the error by about 0.5% and found that models with overlapping pooling generally find it harder to overfit. Thus, while an older architecture, it is still a powerful model capable of achieving high accuracy on very challenging datasets.

D. GoogLeNet Architecture

GoogLeNet is known for reducing error rate to a good extent compared to AlexNet and other state-of-the-art architectures [24]. It is a 22 layers deep network with input image size 224-by-224-by-3. GoogLeNet uses global average pooling as a final layer. These are known as inception modules allowing one to let the network decide whether it wants to pool or convolve. In earlier layers auxiliary classifiers are applied to stabilize the gradient. So, this feed-forwarding idea helps one to figuring out a preliminary classification thus allowing it to bring in the loss at an early stage. So, the deeper network helps to get a vanishing gradient, which is mollified with the auxiliary classifiers. Thus, this feed-forwarding characteristic allows one to decide on the number of inception modules.

The inception module is a combination of multiple 1×1, 3×3 and 5×5 convolution and/or max pooling. These branches are set parallel and then the output is concatenated towards the next layer. This is how the network decides which branches' output could be trusted in the next layer and, thus, which pooling, or convolving is decided.

E. Selecting Pre-trained Model

In this research we use transfer learning by choosing a suitable pre-trained model and training that with our datasets. Here we have used the CASIA iris dataset, versions 1 and

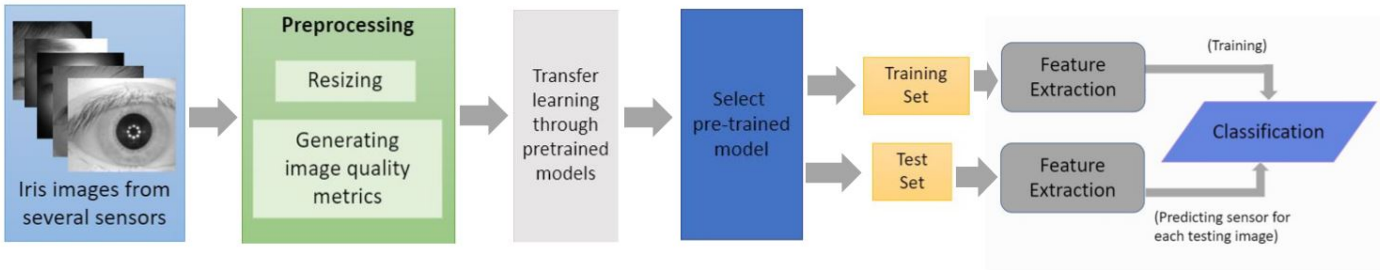


Fig. 3: Our proposed methodology includes transfer learning using iris images of higher quality via the utilization of a set of image quality metrics and achieves close to a hundred percent accuracy after cross-validation.

4, to train the following two networks, namely GoogLeNet and AlexNet, which have reputation in yielding high accuracy classification rates.

GoogLeNet has 4 million parameters compared to AlexNet's that has 60 million. Thus, for the dataset sizes at hand, we empirically experimented which ones perform better with respect to time and accuracy. For this research, we trained both the networks with our data and run a set of experiments. Both networks yield similar accuracy after we apply image quality metrics. However, the former network is much more time consuming. Considering the high amount of time taken by GoogLeNet, we chose AlexNet to be more suitable for this research.

F. Training Through Transfer Learning

We have employed the MATLAB Deep Network Designer and Experiment Manager App to train the AlexNet network for this research. We split the datasets into 90% for training and 10% for testing and repeat the same experiment for different numbers of images (such as 300, 600, 900, 1200 and 1400) every time. The learning rate we chose ranges from 0.001 to 0.007 with a maximum of 10 epochs and trained through a mini batch size of 32. The experiments are defined as a five-class problem where each class represents individual camera sensor accepting near infrared technology, including OKI IRISPASS-h, CASIA close-up camera, CASIA NIR mobile V1 and V2 and a domestic mobile phone.

We choose the training option as Stochastic Gradient Descent with Momentum (SGDM) for optimizing the network. SGDM is known to be very effective for converging very faster by accelerating gradients vectors in the right directions.

IV. Experimental Results

A. Datasets

We evaluate the proposed method by applying it to CASIA IRIS datasets. For this research we used CASIA-Iris-Mobile-V1.0 and CASIA-IrisV4 datasets, which are available online. CASIA IRIS dataset version 1 and 4 includes a total of nine sets of different iris images. CASIA-Iris-Mobile-V1.0 contains a total of 11,000 images from 630 Asian subjects. It includes three subsets based on three different mobile devices identified as S1, S2 and S3. All images were collected under NIR (Near Infrared) illumination and two eyes were captured simultaneously. Images are 8 bit gray-level files stored as JPG format.

CASIA-IrisV4 comprises of six subsets with a total of 54,607 iris images from more than 1,800 genuine subjects and 1,000 virtual subjects. All iris images are 8 bit gray-level JPEG files, collected under near infrared illumination or synthesized. In Table I we do the listing of all five models used in this research with their characteristics such as number, size, features etc. In the Fig. 1 we show the images coming from each individual datasets used. The datasets characteristics are discussed further below.

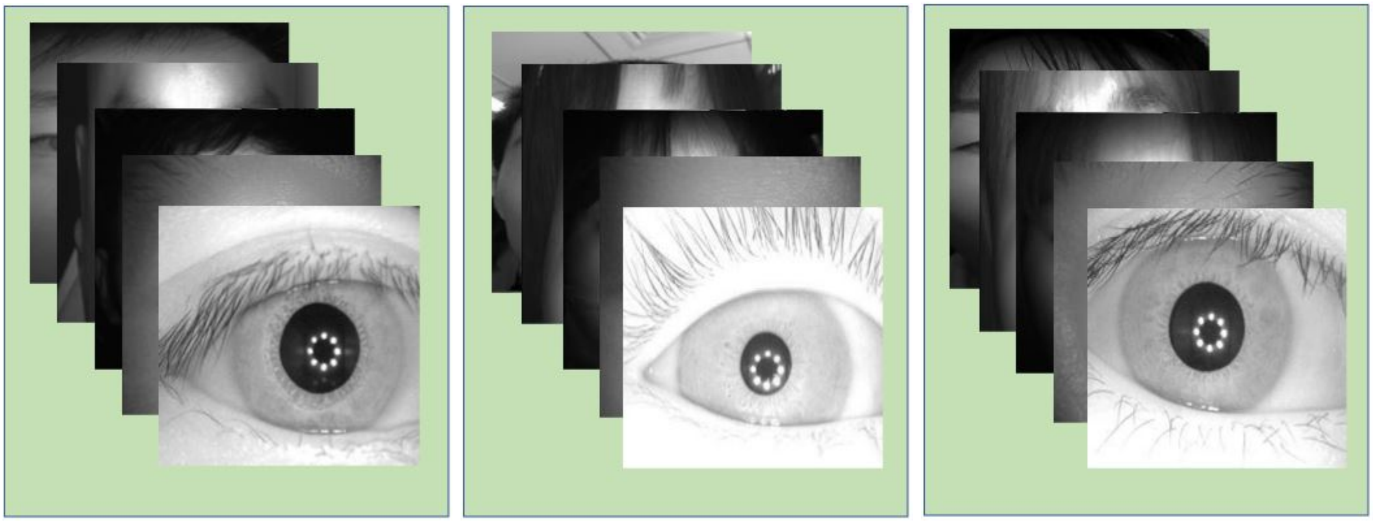
1) CASIA-Iris-M1-S1: Images in this dataset are pictured using the NIR iris imaging module consisting of an NIR camera and multiple NIR illuminator. Its small size (about 5cm×2cm×1cm) makes it easier to attach to a mobile phone through a micro USB port. These images are captured from about 25cm standoff distance and the resolution of the iris images is 1080×1920.

2) CASIA-Iris-M1-S2: Images in this dataset are pictured using an improved NIR imaging module (CASIA NIR mobile module V2). Images captured in this dataset are of 200 Asian subjects and they are collected from three different standoff distances (20cm, 25cm and 30cm, respectively), where 10 images are collected from each distance. The resolution of the iris images is 1968×1024.

3) CASIA-Iris-M1-S3: Images in this dataset are pictured using a domestic mobile phone with NIR iris scanning technology. Images captured in this dataset are of 360 Asian subjects and the resolution of the iris images is 1920×1920.

4) CASIA-Iris-Interval: Images in this dataset are pictured using a close-up iris camera developed by CASIA. They claim that this camera can capture very clear iris images, which makes it very suitable for studying the detailed texture of iris images. The resolution of the iris images is lower than those captured by the other iris sensors discussed above, i.e. 320×280.

5) CASIA-Iris-Lamp: Images in this dataset are pictured using a handheld iris sensor produced by OKI. While capturing these images a lamp was turned on and off. The resolution of



(a) Subject 1 (b) Subject 2 (c) Subject 3

Fig. 4: Different sets of images coming from same subject per set but captured with different sensors

TABLE I: A summary of the characteristics of all iris datasets used in this study.

Subset Characteristics	CASIA-Iris-M1-S1	CASIA-Iris-M1-S2	CASIA-Iris-M1-S3	CASIA-Iris Interval	CASIA-Iris Lamp
Sensor	CASIA NIR mobile module V1	CASIA NIR mobile module V2	A domestic mobile phone with NIR iris-scanning technology	CASIA close-up iris camera	OKI IRISPASS-h
Environment	Indoor	Indoor	Indoor	Indoor	Indoor with lamp on/off
Attributes of Subjects	Most are graduate students of CASIA	Most are workers	Most are students of China	Most are graduate students of CASIA	Most are graduate students of CASIA
No. of Images	1400	6000	3600	2639	16212
Resolution	1080×1920	1968×1024	1920×1920	320×280	640×480
Features	The first NIR mobile iris dataset	Images are collected at three different distances (20, 25, 30 cm)	Images are captured by a mobile phone. The number of subjects is the largest	Cross-session iris images with extremely clear iris texture details	Nonlinear deformation due to variations of visible illumination

the iris images is lower than those captured by the other iris sensors discussed above, i.e. 640×480.

B. Classification and Model Validation

We train the network with MATLAB Experiment Manager and choose the training parameters. We evaluate the result of classification with a loss function graph. A loss function represents an amount of wrong prediction in a classification. In Fig. 5 it is shown that with the increase of epochs the loss function values become lower and the system converges after second epoch.

In the box plot shown in Fig. 7 we show the accuracy at every step. The entire box plot ranges in one point which is 100% and the minimum outlier value is 99.333% for 300 images per class and 20% for 900, 1200 and 1400 images per class. There is no outlier for 600 images per class. Analyzing the box plot, image quality metrics plays a vital role in selecting better quality images, thus making it convenient to perform the experiment with lower number of images per class. It saves both time and computational complexity of the experiments.

The confusion matrix shown in Fig.6 also shows similar characteristics thus ensuring the high accuracy of our proposed method.

Providing the efficiency of the algorithm proposed in this research, we further investigate with different numbers of images per class and with learning rates ranging from 0.001 to 0.007 with the increase of 0.002 at every step. We observe that the number of images per class has much less to do with the accuracy when the learning rate is inclined towards the smallest number selected empirically. However, when the learning rate is 0.005 or 0.007 the accuracy fluctuates. Fig. 8 shows three ROC curves where the area under curve (AUC) is 0.5 when the learning rate is 0.005, whereas both for learning rates 0.003 and 0.001 the AUC is exact 1. Figures shown

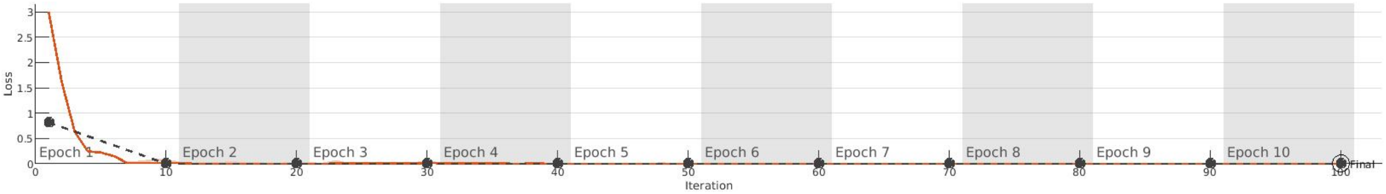


Fig. 5: Loss function showing the loss of accuracy is zero after second epoch.

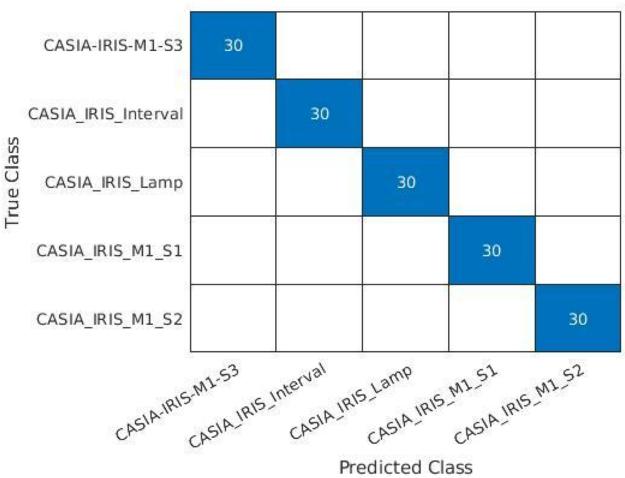


Fig. 6: Confusion matrix shows that the accuracy for sensor identification using AlexNet is 100% for the five datasets used in this research. The Vertical class names here represent true classes and the predicted classes are shown horizontally.

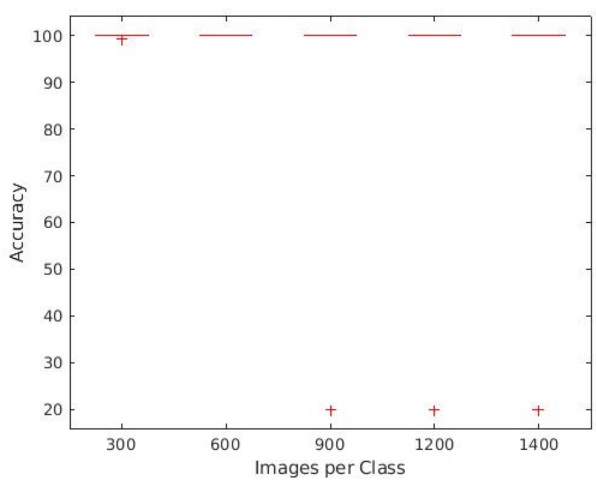


Fig. 7: The box plot above shows the distribution of accuracy for every epochs for different size of datasets used. This shows that our experiment can give better results with smaller amount of data. Although the results for larger datasets are showing some outliers much different than the mean, they occur when learning rate is larger than 0.004.

in here represent the number of images per class as 300. Repeating the experiment for number of images per class as 600, 900, 1200 and 1400 only bolsters this claim.

V. Limitations and Future Work

We have come across some literature [25] mentioning CASIA version 1.0 data set being edited in the pupil area thus referring not to use that in any bio-metric research. Here in this research, we are not doing any human identification for which the above concern might be an issue. The purpose of this research is to solely focus on identifying image sensor, for which editing an image to some extent should not effect the signature of the sensor used to capture the image.

We would like to expand our research to combining multiple biometric data (face, iris, fingerprint) together with different types of sensor data for the same people to investigate the relation between those characteristics and the sensor influence. We also want to elaborate our research with more pre-trained models solely based on mobile phone images as there is a lot of active research and interest capturing and processing mobile phone biometric data these days.

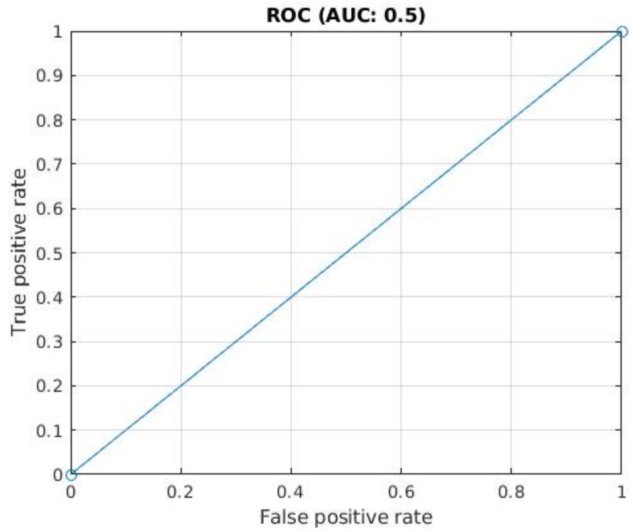
VI. Conclusion

Sensor interoperability is very important in multiple fields, from data organization to forensics [14] and security. This issue might be particularly important for large bio-metric systems, where there must be several sensors involved. Here in this paper we propose a method to choose highest quality images to train popular pre-trained deep learning models.

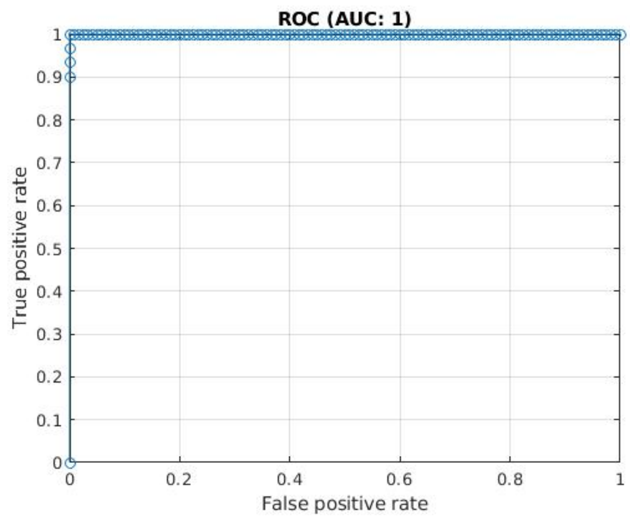
For performing this research, we used datasets from CASIA iris datasets, version 1 and 4. We choose 5 sets of images, find the best quality images from each dataset by applying image quality metrics, trained a transfer learning model and yielded close to 100% accuracy. This methodological model can be extended to more complex biometric research problem related to the need to understand the sensor origin of different biometric data.

Acknowledgment

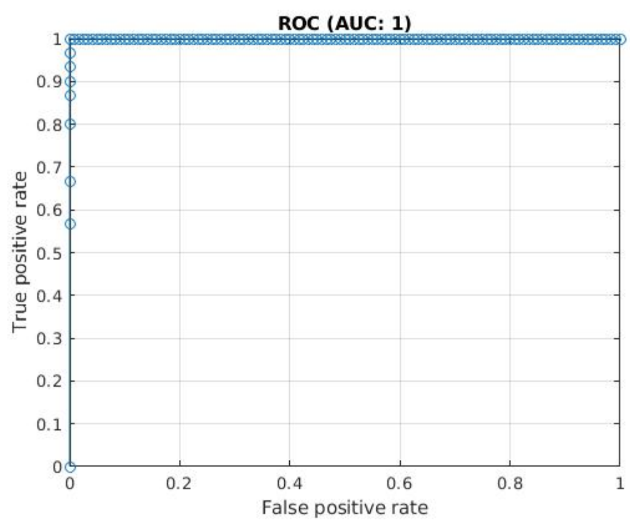
Portions of the research in this paper use the CASIA-Iris-Mobile-V1.0 and CASIA-IrisV4 collected by the Chinese Academy of Sciences' Institute of Automation (CASIA) [9]. We also thank Professor Dr. Natalia Schmid for her valuable suggestions on improving the quality of the paper. The funds to support this registration are provided by UGA via Dr. Bourlai's multispectral imagery lab.



(a) ROC for learning rate 0.005



(b) ROC for learning rate 0.003



(c) ROC for learning rate 0.001

Fig. 8: The AUC in ROC curves above represent the classifier performance indicating that the lower the learning rate, the better the accuracy is.

REFERENCES

[1] F. Marra, G. Poggi, C. Sansone, and L. Verdoliva, "A deep learning approach for iris sensor model identification," *Pattern Recognition Letters*, vol. 113, pp. 46–53, 2018.

[2] S. S. Arora, M. Vatsa, R. Singh, and A. Jain, "On iris camera interoperability," in *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 2012, pp. 346–352.

[3] K. W. Bowyer, S. E. Baker, A. Hentz, K. Hollingsworth, T. Peters, and P. J. Flynn, "Factors that degrade the match distribution in iris biometrics," *Identity in the information Society*, vol. 2, no. 3, pp. 327–343, 2009.

[4] A. Ross and A. Jain, "Biometric sensor interoperability: A case study in fingerprints," in *International Workshop on Biometric Authentication*. Springer, 2004, pp. 134–145.

[5] A. Ross and R. Nadgir, "A calibration model for fingerprint sensor interoperability," in *Biometric Technology for Human Identification III*, vol. 6202. International Society for Optics and Photonics, 2006, p. 62020B.

[6] G. R. Doddington, M. A. Przybocki, A. F. Martin, and D. A. Reynolds, "The NIST speaker recognition evaluation–overview, methodology, systems, results, perspective," *Speech communication*, vol. 31, no. 2-3, pp. 225–254, 2000.

[7] N. Bartlow, N. Kalka, B. Cukic, and A. Ross, "Identifying sensors from fingerprint images," in *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 2009, pp. 78–84.

[8] A. Ross and R. Nadgir, "A thin-plate spline calibration model for fingerprint sensor interoperability," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 8, pp. 1097–1110, 2008.

[9] Chinese Academy of Sciences' Institute of Automation (CASIA). CASIA-Iris-Mobile-V1.0. (2008, Dec 27). [Online]. Available: http://biometrics.idealtest.org/

[10] H. Cao and A. C. Kot, "Accurate detection of demosaicing regularity for digital image forensics," *IEEE Transactions on Information Forensics and Security*, vol. 4, no. 4, pp. 899–910, 2009.

[11] K. San Choi, E. Y. Lam, and K. K. Wong, "Source camera identification by jpeg compression statistics for image forensics," in *TENCON 2006-2006 IEEE Region 10 Conference*. IEEE, 2006, pp. 1–4.

[12] L. T. Van, S. Emmanuel, and M. S. Kankanhalli, "Identifying source cell phone using chromatic aberration," in *2007 IEEE International Conference on Multimedia and Expo*. IEEE, 2007, pp. 883–886.

[13] C. Galdi, M. Nappi, and J.-L. Dugelay, "Multimodal authentication on smartphones: Combining iris and sensor recognition for a double check of user identity," *Pattern Recognition Letters*, vol. 82, pp. 144–153, 2016.

[14] A. T. Ho and S. Li, "Forensic camera model identification," 2015.

[15] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.

[16] A. Uhl and Y. Höller, "Iris-sensor authentication using camera prnu fingerprints," in *2012 5th IAPR International Conference on Biometrics (ICB)*. IEEE, 2012, pp. 230–237.

[17] O. Çeliktutan, B. Sankur, and I. Avcibas, "Blind identification of source cell-phone model." *IEEE Trans. Information Forensics and Security*, vol. 3, no. 3, pp. 553–566, 2008.

[18] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[19] T. Gloe, "Feature-based forensic camera model identification," in *Transactions on data hiding and multimedia security VIII*. Springer, 2012, pp. 42–62.

[20] G. Xu and Y. Q. Shi, "Camera model identification using local binary patterns," in *2012 IEEE International Conference on Multimedia and Expo*. IEEE, 2012, pp. 392–397.

[21] A. Tuama, F. Comby, and M. Chaumont, "Camera model identification based machine learning approach with high order statistics features," in *2016 24th European Signal Processing Conference (EUSIPCO)*. IEEE, 2016, pp. 1183–1187.

[22] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep learning for steganalysis via convolutional neural networks," in *Media Watermarking, Security, and Forensics 2015*, vol. 9409. International Society for Optics and Photonics, 2015, p. 94090J.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[24] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[25] P. J. Phillips, K. W. Bowyer, and P. J. Flynn, "Comments on the casia version 1.0 iris data set," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 10, pp. 1869–1870, 2007.