

Off-Angle Iris Recognition using Single and Multiple Deep Learning Frameworks

David Chavarro and Mahmut Karakaya

Dept of Computer Science, Kennesaw State University, Marietta, GA USA 30060

ABSTRACT

Iris recognition is a widely used biometric technology that has high accuracy and reliability in well-controlled environments. However, the recognition accuracy can significantly degrade in non-ideal scenarios, such as off-angle iris images. To address these challenges, deep learning frameworks have been proposed to identify subjects through their off-angle iris images. Traditional CNN-based iris recognition systems train a single deep network using multiple off-angle iris image of the same subject to extract the gaze invariant features and test incoming off-angle images with this single network to classify it into same subject class. In another approach, multiple shallow networks are trained for each gaze angle that will be the experts for specific gaze angles. When testing an off-angle iris image, we first estimate the gaze angle and feed the probe image to its corresponding network for recognition.

In this paper, we present an analysis of the performance of both single and multimodal deep learning frameworks to identify subjects through their off-angle iris images. Specifically, we compare the performance of a single AlexNet with multiple SqueezeNet models. SqueezeNet is a variation of the AlexNet that uses 50x fewer parameters and is optimized for devices with limited computational resources. Multi-model approach using multiple shallow networks, where each network is an expert for a specific gaze angle. Our experiments are conducted on an off-angle iris dataset consisting of 100 subjects captured at 10-degree intervals between -50 to +50 degrees. The results indicate that angles that are more distant from the trained angles have lower model accuracy than the angles that are closer to the trained gaze angle. Our findings suggest that the use of SqueezeNet, which requires fewer parameters than AlexNet, can enable iris recognition on devices with limited computational resources while maintaining accuracy. Overall, the results of this study can contribute to the development of more robust iris recognition systems that can perform well in non-ideal scenarios.

Keywords: biometrics, iris recognition, off-angle iris images, deep learning.

1. INTRODUCTION

Biometric recognition has found wide-ranging applications, including security systems, access control, and surveillance. Among the various biometric modalities, iris recognition stands out for its remarkable accuracy and dependability in identifying individuals. Typically, iris recognition systems acquire images of the iris, which is a unique and stable biometric characteristic, and utilize computer algorithms to extract and compare features to facilitate identification. For example, at certain airports, passport control officers use iris recognition cameras to verify the passengers' credentials [1]. However, this requires the iris recognition system to be operated under limited conditions.

The traditional iris recognition method involves several steps to identify a person based on their iris pattern. First, a digital camera captures an image of the iris, which is then isolated from the rest of the image using image processing techniques. The iris is then normalized to correct for any distortion caused by the camera or the viewing angle. Next, unique features of the iris, such as the texture of the iris stroma and the radial pattern of the iris furrows, are extracted and used to create a template that represents the iris. The template is then compared to a database of stored templates to identify the person. Finally, a decision is made about the identity of the person based on the similarity between the templates, using a threshold to determine whether the new iris image matches any of the stored templates. The traditional iris recognition setup only allows for accurate identification when the subjects are facing the camera. Despite its high accuracy in controlled environments, iris recognition can be significantly affected by non-ideal scenarios, such as off-angle iris images [2]. While

* Email: mkarakay@kennesaw.edu; Phone: +1(470) 578 6983.

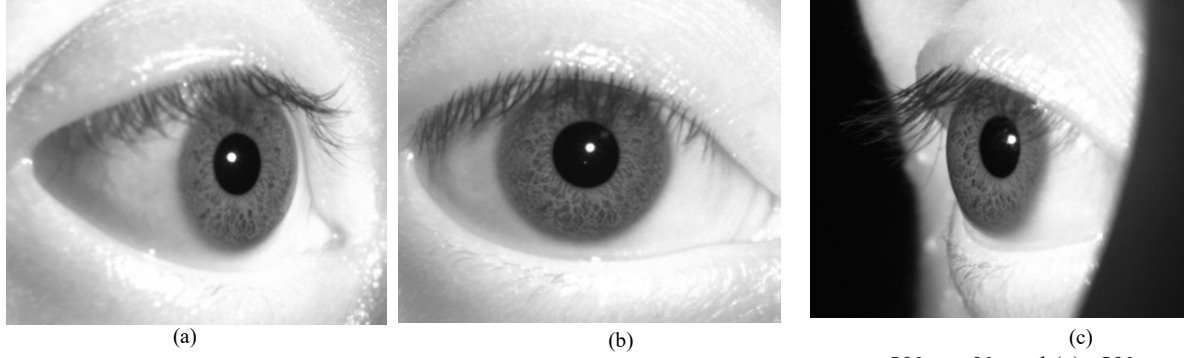


Figure 1: Sample frontal and off-angle iris images captured at different angles (a) $+50^\circ$, (b) 0° , and (c) -50° .

it is effective in identifying subjects with less than 20° gaze difference, its accuracy reduced at steeper gaze angles, possibly due to the difference in the elliptical shape of the irises between frontal and non-frontal gaze angles [3].

Off-angle iris images pose a significant challenge to iris recognition as the angle of deviation from the frontal view can introduce inconsistencies in the iris texture, resulting in reduced recognition accuracy as shown in Figure 1. To address this issue, various approaches have been proposed and investigated in biometrics research. One approach involves using 3D reconstruction of the iris to estimate its orientation and align it to a frontal view. Another approach uses generative models to synthesize frontal view iris images from off-angle views, trained on a large dataset of frontal view iris images using deep neural networks. Alternatively, a multi-view representation of the iris can be employed, where images of the iris are captured from multiple angles and combined into a single representation. This technique can match off-angle iris images to frontal view templates using a combination of similarity scores from different angles. Overall, recognizing off-angle iris images is an active and challenging area of research in biometrics, and various techniques are being developed to enhance recognition accuracy and reliability.

Recent research has introduced deep learning frameworks as a potential solution to the challenge of recognizing off-angle iris images, especially Convolutional Neural Networks (CNNs). CNNs are specifically designed for image classification tasks and can be trained on a large dataset of labeled iris images to classify new images into specific individuals [4]. During training, the CNN learns to extract relevant features from the iris images using convolutional and pooling layers, and the learned features are then passed through fully connected layers to produce the final classification output. CNN-based iris recognition frameworks can learn discriminative features that are specific to individual irises, even under challenging conditions such as occlusion and variations in lighting. Additionally, CNNs can be trained to recognize off-angle iris images, which traditional iris recognition methods struggle with. Studies have shown high recognition accuracies using CNNs for iris recognition, with some achieving state-of-the-art performance on benchmark datasets. However, further improvements are necessary to reduce computational costs and increase robustness to noisy and low-quality images.

This paper presents an investigation of the performance of single and multimodal deep learning frameworks for the identification of subjects through their off-angle iris images. In traditional CNN-based iris recognition systems, a single deep network is trained using multiple off-angle iris images of the same subject to extract gaze-invariant features. However, this approach may not be adequate for recognizing off-angle images that differ significantly from the training set. To address this limitation, multiple shallow networks are trained for each gaze angle, which function as experts for specific gaze angles. During testing, an off-angle iris image is used to estimate the gaze angle, which is then used to identify the corresponding network for recognition. The contribution of this paper is to analyze the effects of gaze angle on CNN-based off-angle iris recognition and compare the recognition performance of single and multimodal CNN models.

The remainder of this paper is structured as follows. Section 2 offers a comprehensive review of previous research in off-angle iris recognition, including a description of the traditional methodology used for identifying subjects and an explanation of the architecture of traditional CNNs. Section 3 outlines our proposed single-modal and multi-modal CNN models, describing their training pipelines and the composition of their neural network layers. In Section 4, we present and analyze the results of our experiments. Finally, Section 5 provides a conclusion and suggestions for further research.

2. RELATED WORKS

Traditional iris recognition systems have been developed to authenticate the identity of a subject. The methodology for these systems was that it required the user to look at the iris camera. The traditional iris recognition setup only allows for accurate identification when the subjects are facing the camera. While this was effective in correctly identifying the subject, it was inconvenient to subjects because it required them to stand still. Despite its high accuracy in controlled environments, iris recognition can be significantly affected by non-ideal scenarios, such as off-angle iris images. Off-angle iris recognition is a challenging problem in biometrics, and there have been several approaches proposed in the literature to address this issue. This section reviews the most relevant works on off-angle iris recognition using deep learning frameworks.

Advancements in the field of neural networks have led to the development and application of complex deep learning architectures, including Convolutional Neural Networks (CNNs), in various research areas and practical uses. An initial work in utilizing a deep learning framework to extract iris texture features for recognition was DeepIrisNet [5]. Unlike traditional iris recognition systems that utilize Gabor filters to transform iris texture to binary iris codes, DeepIrisNet extracts features from a normalized iris image, with matching performed in the final fully connected layer by comparing Euclidean distances between subjects. Liu et al. [6] proposed a CNN with nine layers that enhances recognition performance by learning pairwise filters from diverse sources. To improve periocular biometrics in visible spectrum images, Proenca and Neves [7] proposed deep-PRWIS using a CNN framework based on AlexNet, which excludes the iris and sclera from the periocular region through artificial sampling to improve recognition performance given the various degradation factors for these ocular structures present in visible spectrum images. Zhao and Kumar [8] presented a study to enhance recognition performance for the periocular region by proposing a semantics-assisted CNN that trains multiple deep networks to learn additional semantic information from the periocular region, such as ethnicity and gender, with recognition performance improved through combining extra information and output features from the network. While these studies have improved recognition performance for ocular and periocular structures, they focus solely on frontal angle images and do not address challenges associated with identifying off-angle images.

This study builds upon our previous work in [9], which tackled the issue of performance degradation in off-angle images resulting from differences in gaze angle. In our previous study, we presented deep learning frameworks based on convolutional neural networks (CNNs) to enhance the recognition accuracy of off-angle iris, ocular, and periocular biometric modalities. We compared the performance of popular CNN architectures such as AlexNet, GoogLeNet, and ResNet50 for off-angle biometrics. Our research demonstrated that including the ocular and periocular region along with the iris resulted in improved recognition results for off-angle images. To achieve this, we trained a single deep network using multiple off-angle iris images of the same subject to extract gaze-invariant features. However, the large size of this network may not be feasible for devices with limited computational resources. To address this, we proposed a multi-model approach using multiple shallow networks, where each network specializes in recognizing iris images at a specific gaze angle. So, we can use smaller, lightweight shallow networks for off-angle iris recognition while achieving high accuracy.

3. METHODOLOGY

This section describes the proposed approach for off-angle iris recognition using both single-modal and multi-modal CNN models and we compare the performance of two CNN models: AlexNet and SqueezeNet. We first briefly describe AlexNet and SqueezeNet. Second, we will introduce the transfer learning technique to employ the pre-trained models. Then, we will introduce the details of single and multi-model deep network frameworks. We present an analysis of the performance of both single and multimodal deep learning frameworks to identify subjects through their off-angle iris images. Specifically, we compare the performance of a single AlexNet with multiple SqueezeNet models.

The AlexNet [10] is a convolutional neural network (CNN) model comprised of a combination of layers to enable object identification. The model consists of five convolutional layers, multiple max-pooling layers, and three fully connected layers where it contains 61 million parameters. Its size in MATLAB is 227 MB. Each layer in the network applies various filters of different sizes to the outputs of the preceding layer, which is then passed through a rectified linear unit (ReLU) activation function and a max-pooling layer. The ReLU function is a simple and non-linear function that sets all negative input values to zero, allowing positive values to pass through unaltered. The max-pooling layer reduces the feature map's spatial dimensions while preserving critical features. Additionally, the model employs dropout layers to prevent overfitting during training by randomly blocking certain features and setting them to zero when passing to the next layer. The final convolutional layer's output is flattened and fed into the fully connected layers, ultimately outputting class probabilities using a softmax function.

SqueezeNet is a convolutional neural network (CNN) architecture that is optimized for devices with limited computational resources. Compared to traditional CNN models like AlexNet, SqueezeNet has a significantly smaller number of parameters (1.24 million), which makes it more efficient in terms of memory usage and processing speed [11]. Its size in MATLAB is 5.2 MB. SqueezeNet uses Fire modules, which consist of a combination of 1x1 and 3x3 convolutional filters. The 1x1 filters are used in the squeeze layer, which reduces the number of input channels to the following expand layer by 9x. The expand layer then uses a combination of 1x1 and 3x3 filters to increase the number of output channels. SqueezeNet also uses pooling layers between certain Fire modules to ensure that the activation map size is at least 1x1. The ReLU activation function is used for Fire and squeeze layers. Additionally, SqueezeNet also uses a concept called bypass connections to preserve information from earlier layers and prevent information loss during training.

The AlexNet and SqueezeNet are trained on the ImageNet dataset that has more than 1.2 million labeled images in 1,000 different classes. Due to the limited number of iris images for each subject, we use the transfer learning technique instead of training deep models from scratch. Transfer learning allows for the leveraging of knowledge and expertise gained from a pre-trained model by using its pre-trained weights, which were obtained by training the model on large datasets. To employ transfer learning, the last fully connected layer is initially replaced with a new fully-connected layer that has the same number of neurons as the number of subjects in the iris dataset. Finally, the model is trained using the iris dataset.

Traditional CNN-based iris recognition systems utilize a single deep neural network to extract gaze-invariant features for recognizing iris patterns. These systems are trained on multiple off-angle iris images of the same subject to ensure that the extracted features are invariant to gaze angle. During the training process, the single-modal network such as AlexNet is trained with iris images from all gaze angles to enable it to classify the subject's identity from any given test image captured at any gaze angle. This approach aims to create a robust iris recognition system that is not affected by variations in the gaze angle. However, it is required to use a deeper network to learn the texture variations in the off-angle iris images compared with the only frontal iris images. As an alternative approach, we propose using multiple shallow networks such as SqueezeNet that are trained for each gaze angle that serve as experts for specific gaze angles. When testing an off-angle iris image, we first estimate the gaze angle and feed the probe image to its corresponding network for recognition.

The first step for the multi-modal approach is gaze estimation to predict the gaze angle of the subject given an input image. In our previous study [12], we developed a CNN-based gaze estimation model using pre-trained AlexNet model to estimate the gaze angle of iris image with 94% accuracy. The model was re-trained via transfer learning to satisfy this use case's requirements. Five convolutional layers were used, along with three fully connected layers and a softmax layer. The training dataset of the original experiment used was the same dataset that we are using for this experiment. As a second step in multi-modal approach, we trained multiple shallow SqueezeNet networks for each gaze angle. After the gaze angle estimation, we feed the iris image to corresponding CNN-based iris recognition network and use their classification result.

4. EXPERIMENTAL SETUP AND RESULTS

Our study was conducted using an off-angle iris dataset that includes approximately 10,000 images from around 100 subjects captured by an IDS-UI-3240ML-NIR camera sensitive to near-infrared light. These images were of both the left and right eyes of the subjects and were taken at angles ranging from -50° to 50° with 10° increments. The images were captured by a horizontally moving camera to capture all angles, capturing 10 iris images per step, resulting in a total of 10 frontal and 100 off-angle iris images per subject. Example images from each angle can be viewed in Fig. 2. Additional details regarding our off-angle iris dataset are available in [13].

We employed the MATLAB integrated development environment and an Nvidia GeForce RTX 2080 SUPER with a 16GB memory to train the SqueezeNet, AlexNet, and Resnet50 CNN models. The initial step in the training process involves processing the images in the dataset using the image datastore library. This library enables the loading of each image dynamically from the dataset directory and parsing of the image attributes based on the file naming conventions. Subsequently, the training and validation datasets are created for the CNN models. The pre-trained model is used to train the single-modal and multi-modal CNN models during the training process. In the experiments, the models are customized by replacing the fully connected layer, SoftMax layer, and classification layer with the remaining three layers. For the AlexNet models, the learning drop rate, learning drop period, initial learning rate, max epoch, and minibatch size are set to 0.2, 5, 10-4, 60, and 6, respectively. We use the stochastic gradient descent (SGDM) optimizer. For the single-modal and multi-modal SqueezeNet models, the convolutional layer conv10 has a [1,1] stride with the output number matching the number of subjects. Moreover, the classification layer is replaced, which employs the output of conv10.

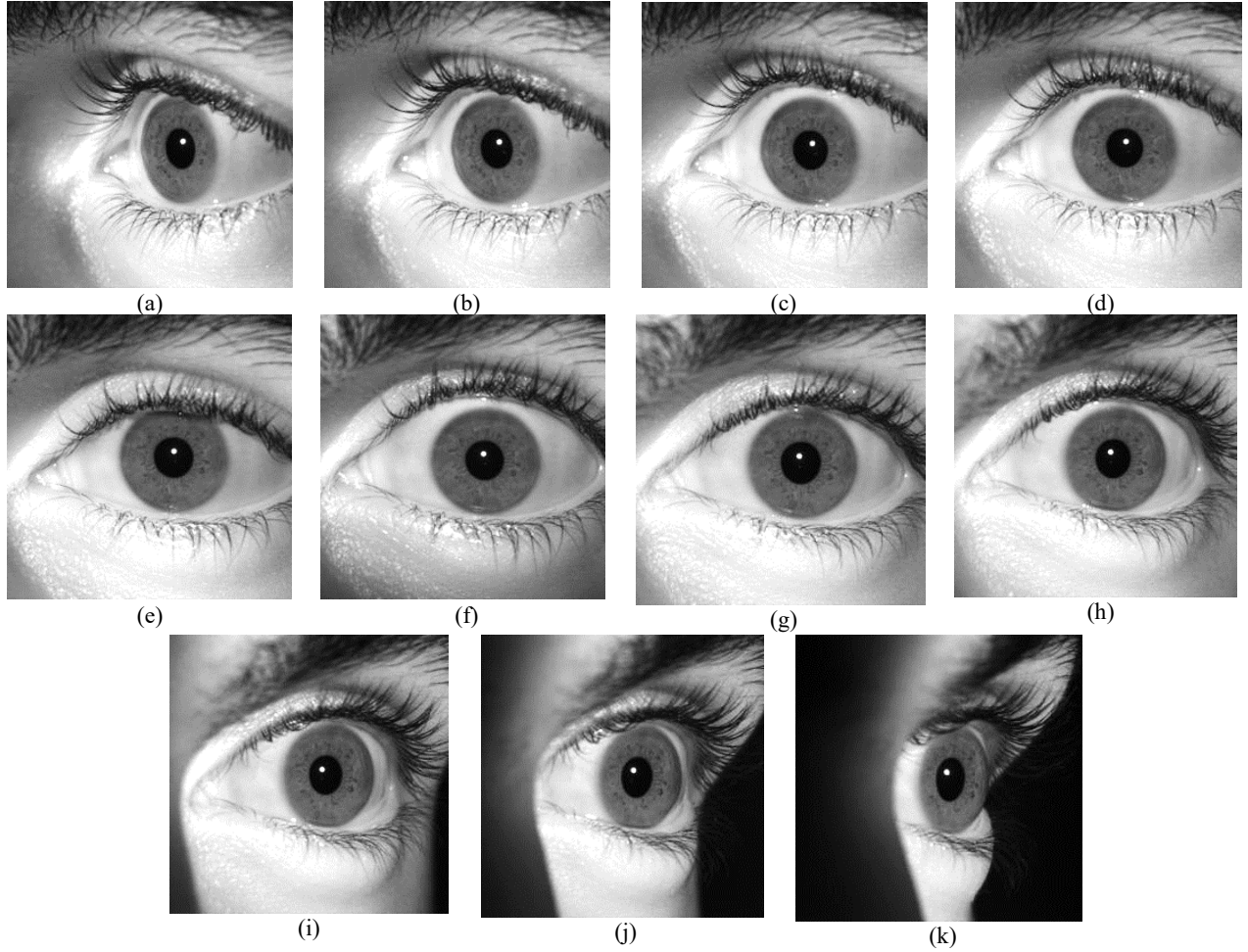


Figure 2. Example images from off-angle iris databases, (a)-50° angle, (b)-40° angle, (c)-30° angle, (d)-20° angle, (e)-10° angle, (f) 0° angle, (g) 10° angle, (h) 20° angle, (i) 30° angle, (j) 40° angle, (k) 50° angle.

To train the single-modal model, the training dataset was chosen as a subset image of the original dataset from all training gaze angles, where one image per gaze angle was used from each subject. The remaining images of the dataset became part of the validation dataset. For the multi-modal models that classified subjects, the training dataset was a subset of images that were taken at a specific gaze angle, while the validation dataset was made up of the remaining images not used for training. To evaluate the performance of the single-modal models, the accuracy for each gaze angle is calculated. We also evaluated the recognition performance using the ROC curves.

In our first set of experiments, we evaluated the performance of SqueezeNet and AlexNet with frontal training and off-angle testing. To train both models, only frontal iris images were used for training and other gaze angles were used for the test to assess the trained models. Fig 3 shows the classification accuracy of AlexNet and SqueezeNet models for iris images at different gaze angles. The x-axis represents the off-angle gaze angle in degrees, ranging from -50 to 50 degrees, while the y-axis represents the classification accuracy of the models. As shown in Fig. 3, the results suggest that both models are capable of accurately classifying subjects based on their frontal iris images, but their performance decreases as their gaze angles increase. Based on the results, AlexNet outperforms SqueezeNet especially for off-angle iris images when networks trained with frontal images. The main reason is that AlexNet is a deeper model with more layers and more trainable parameters than SqueezeNet. It allows AlexNet to learn more complex features and patterns from the input data. In contrast, SqueezeNet is designed to be a more lightweight model with a smaller memory footprint, and its use of smaller convolutional kernels may limit its ability to learn complex patterns in the input data.

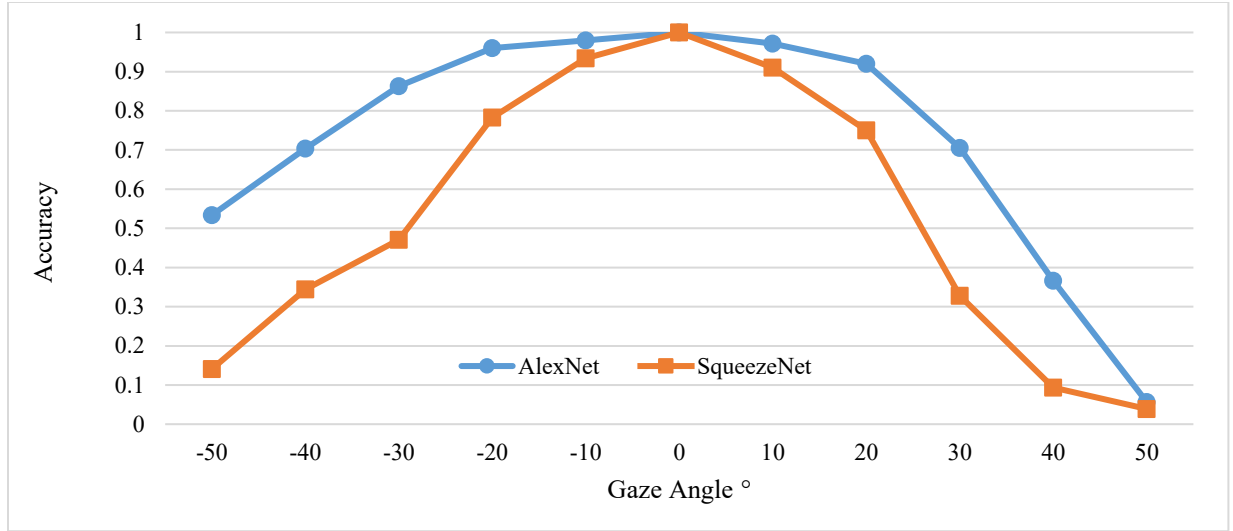


Figure 3. The accuracies of both AlexNet and SqueezeNet per gaze angle.

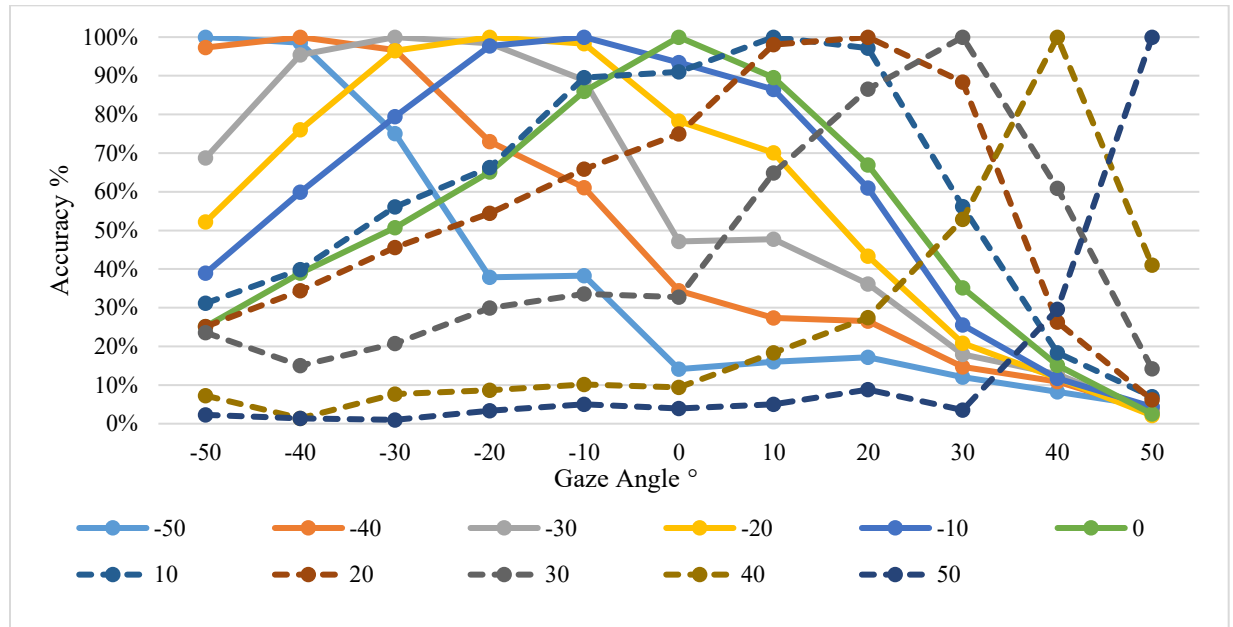


Figure 4. The accuracies of the multi-modal SqueezeNet models. all training images per gaze angle

In the second set of experiments, we aimed to investigate the recognition performance of SqueezeNet when trained with images from off-angle views and tested with other off-angle views. Each experiment involved training the model with images from only one gaze angle between -50 and 50 degrees. Fig. 4 displays the recognition accuracy of SqueezeNet trained with a subset of images captured at different gaze angles and evaluated on images from both the trained gaze angles and additional off-angle gaze angles. Each line in the graph represents the accuracy of the model that was trained at a specific gaze angle. The x-axis and y-axis represent the trained and evaluated gaze angles, respectively, with values ranging from -50 to 50 degrees.

Our results showed that as the evaluated gaze angle deviated further from the trained gaze angle, the classification accuracy decreased. Specifically, when the trained gaze angle was -50 degrees and the evaluated angle was 50 degrees, the accuracy dropped to 2%. graph demonstrates that the performance of the model is affected by both the trained gaze angle and the

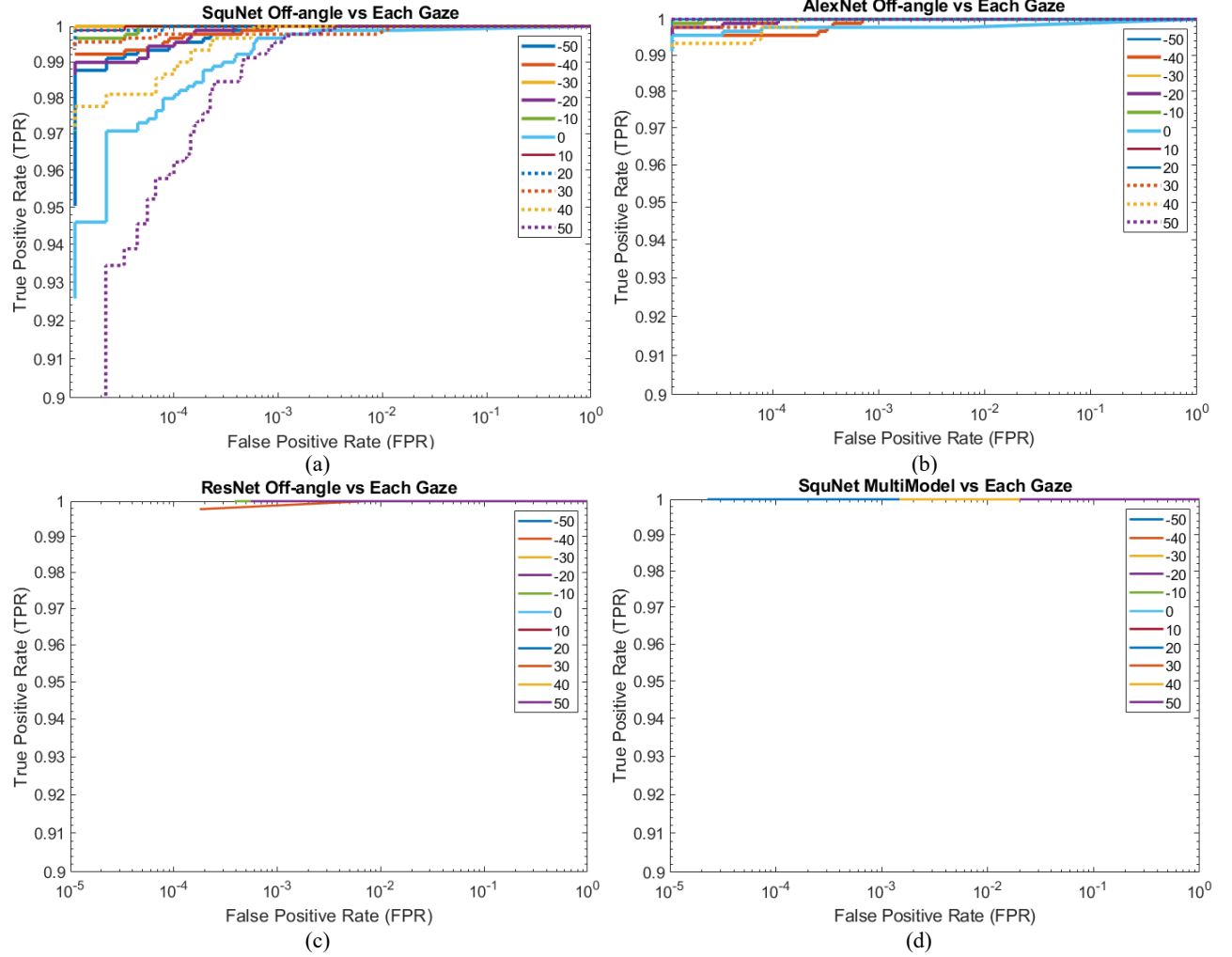


Figure 5. Receiver Operator Characteristic (ROC) curves for (a) SqueezeNet, (b) AlexNet, and (c) ResNet50 where a single model trained with multiple off-angle iris images from different gaze angles. (d) Multi-model SqueezeNet.

evaluated gaze angle. Generally, the accuracy was higher when the evaluated angle was closer to the trained angle, and the accuracy decreased as the deviation between the two angles increased. This is due to the difference in the appearance of the iris region in the input images. The iris texture appears differently when viewed from different angles, and when the system is trained on images captured from a specific angle, it learns to recognize iris features from that viewpoint. Thus, when the system is tested on images captured from different angles, it may fail to recognize iris features accurately, leading to a decrease in performance. Moreover, the model achieved better accuracy when evaluating angles that were closer to the trained gaze angle. These results showed that multi-model SqueezeNet may show high performance when off-angle iris images tested with their corresponding model with similar gaze angle.

In the third set of experiments, we investigated the recognition performance of a single model trained with multiple off-angle iris images captured from different gaze angles. To assess the impact of network depth on recognition performance, we utilized three different network architectures: SqueezeNet, AlexNet, and ResNet50. Fig. 5(a-c) displays the receiver operator characteristic (ROC) curves for these models, illustrating the true positive rate (TPR) and false positive rate (FPR) for various threshold values. A ROC curve closer to the upper left corner of the graph indicates better performance. Our results reveal that SqueezeNet exhibits the lowest recognition performance, while ResNet50 demonstrates the highest recognition performance due to its deeper network architecture. AlexNet performs better than SqueezeNet but not as well as ResNet50. These findings suggest that a deep network is better equipped to learn the differences in the appearance of the iris region in off-angle images, which is essential for accurate recognition. Overall, the ROC curves provide a

quantitative and visual evaluation of the recognition performance of the different network architectures, emphasizing the significance of network depth in achieving accurate recognition of off-angle iris images using single model.

As an alternative approach to using a single deep network for off-angle iris recognition, we propose using multiple shallow networks such as SqueezeNet, each trained for a specific gaze angle, serving as experts for recognizing images from that angle. When presented with an off-angle iris image, we first estimate the gaze angle and then feed the probe image to the corresponding expert network for recognition. To evaluate the recognition performance of this approach, we trained 11 SqueezeNet networks, each acting as an expert for a specific gaze angle. Fig. 5(d) displays the recognition performance analysis of the Multi-model SqueezeNet using an ROC plot. Our results showed similar recognition performance for off-angle iris images compared to using a single deep network trained on multiple off-angle iris images. This highlights the potential for using multiple shallow networks as a viable alternative to a single deep network for off-angle iris recognition.

5. CONCLUSION

In this study, we investigated the recognition performance of deep learning models for off-angle iris images. We performed three sets of experiments using SqueezeNet, AlexNet, and ResNet50 to evaluate the impact of network depth on recognition performance. Our results showed that a deeper network such as ResNet50 achieved the best recognition performance, while a shallow network such as SqueezeNet performed the worst. We also observed that the recognition performance of a model decreased as the evaluated gaze angle deviated further from the trained gaze angle. To address this issue, we proposed a multi-model approach using multiple shallow networks, where each network is an expert for a specific gaze angle. Overall, our study highlights the importance of network depth and gaze angle variation in off-angle iris recognition. Our findings suggest that a deep network is better suited for learning the differences in the appearance of the iris region in off-angle images, while a shallow network can serve as an expert for a specific gaze angle. Our proposed multi-model approach can be an effective solution for off-angle iris recognition, offering similar performance compared to a single deep network. Future research can explore the use of other deep learning models and more sophisticated network architectures to further improve the recognition performance of off-angle iris images.

ACKNOWLEDGMENTS

This project was made possible by support from Secure and Trustworthy Cyberspace (SaTC) program of The National Science Foundation (NSF) under grant award CNS-1909276 and CNS-2100483.

REFERENCES

- [1] Wang K. and Kumar, A. "Toward More Accurate Iris Recognition Using Dilated Residual Features," in IEEE Transactions on Information Forensics and Security, vol. 14, no. 12, pp. 3233-3245, (2019).
- [2] Karakaya, M. "A study of how gaze angle affects the performance of iris recognition" Pattern Recognition Letters, 82(2), pp. 132-143, (2016).
- [3] Jalilian, E., Karakaya, M. and Uhl, A. "CNN-Based off-Angle Iris Segmentation and Recognition." IET Biometrics 10, no. 5 pp. 518-35, (2021).
- [4] Tajbakhsh N. et al., "Convolutional neural networks for medical image analysis: full training or fine tuning?" IEEE Trans. Med. Imaging 35(5), 1299-1312 (2016).
- [5] Gangwar A. and Joshi, A. "DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition," In Proc. of the IEEE International Conference on Image Processing, pp. 2301-2305, (2016).
- [6] Liu, N. Zhang, M. Li, H. Sun, Z. and Tan, T. "Deepiris: Learning pairwise filter bank for heterogeneous iris verification," Pattern Recognition Letters, vol. 82, no. 2, pp. 154-161, (2015).
- [7] Proenca H. and Neves, J. "Deep-PRWIS: Periocular Recognition Without the Iris and Sclera Using Deep Learning Frameworks," IEEE Transactions on Information Forensics and Security, 13(4), (2018).
- [8] Zhao Z. and Kumar, A. "Accurate periocular recognition under less constrained environment using semantics-assisted convolutional neural network," IEEE Trans. Inf. Forensics Security, vol. 12, no. 5, (2016).
- [9] Karakaya, M. "Iris-ocular-periocular: toward more accurate biometrics for off-angle images," Journal of Electronic Imaging 30(3) 033035, (2021).
- [10] Krizhevsky, A. Sutskever, I. and Hinton, G. "Imagenet classification with deep convolutional neural networks," In Proceedings of the Advances in Neural Information Processing Systems Conference, pp. 1097-1105, (2012).

- [11] Iandola, F., Han, S., Moskewicz, M., Ashraf, K., Dally, W. and Keutzer, K. "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," In Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pp. 1-13, (2016).
- [12] Diab, K., and Karakaya, M. "CNN-Based Gaze Estimation for Off-angle Iris Recognition," In proceedings of SoutheastCon, pp. 736-742, (2022).
- [13] Karakaya, M. "Deep Learning Frameworks for Off-Angle Iris Recognition," In Proceedings of IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), pp.1-8, (2018).