

Segmentação de Imagens de Íris Utilizando Deep Learning

Diego Filipe Souza de Lima
Universidade Federal da Paraíba, BRA
Email: diego.filipe@gmail.com

Leonardo Vidal Batista
Universidade Federal da Paraíba, BRA
Email: leonardo@ci.ufpb.br

Abstract—Current biometric systems can recognize individuals through various traits such as fingerprint, face, iris, palm, etc. Among these varied characteristics, the iris is one that most needs the collaboration of the individual. On the other hand, it is one of the most reliable forms of recognition because of the unique patterns it has in its composition. However, the use of this trait in a non-cooperative way means that the current systems perform below satisfactory, mainly due to the difficulty of locating and segmenting the iris region, which generates errors that are propagated throughout the recognition process, affecting the final performance of the systems directly. The present work proposes an iris segmentation algorithm using a Deep Learning technique known as Convolutional Autoencoder, which can perform satisfactorily in both cooperative and non-cooperative environments. The satisfactory performance of the algorithm was evident when compared to algorithms present in the literature, using images with similar capture patterns. The results of the segmentation process were evaluated using iris segmentation error and computational vision metrics, then compared with some of the best results found in the literature. The proposed method achieved in some cases an error rate 68% lower than the other algorithms.

Resumo—Sistemas biométricos atuais são capazes de reconhecer indivíduos por intermédio de diversos traços, como: impressão digital, face, íris, palma da mão, etc. Dentre essas variadas características, a íris é uma das que mais necessita da colaboração do indivíduo. Por outro lado, ela é uma das formas mais confiáveis de reconhecimento devido aos padrões extremamente singulares que ela possui em sua composição. Porém, a utilização dessa característica de forma não cooperativa faz com que os sistemas atuais tenham baixo desempenho, principalmente pela dificuldade de localizar e segmentar a região da íris. Isso acarreta erros que são propagados por todo o processo de reconhecimento, afetando diretamente o desempenho final dos sistemas. O presente trabalho propõe um algoritmo de segmentação de íris utilizando uma técnica de Deep Learning conhecida como *Autoencoder* Convolutacional, com desempenho satisfatório tanto em ambientes cooperativos como não-cooperativos. O desempenho superior do algoritmo ficou evidente quando comparado a algoritmos presentes na literatura, utilizando imagens com padrão de captura semelhante. Os resultados do processo de segmentação foram comparados com alguns dos melhores resultados encontrados na literatura e avaliados utilizando métricas de erro de segmentação de íris e de visão computacional. O método proposto foi capaz de alcançar em alguns casos uma taxa de erro de 68% menor do que os outros algoritmos.

I. INTRODUÇÃO

Reconhecimento biométrico por intermédio de características presentes na íris é até o presente momento uma das formas

mais eficientes e confiáveis, uma vez que essa região do olho humano apresenta características únicas e, em geral, um baixo grau de degradação com o passar do tempo [1]. Devido ao fato das características serem extraídas de uma região específica, há uma relação direta dos resultados dos estágios de localização e segmentação, com os resultados do estágio de reconhecimento, tornando assim essencial que esses estágios iniciais apresentem resultados de alta qualidade, para que o sistema de reconhecimento como um todo possa ter uma alta qualidade [2], [3].

Os algoritmos convencionais de reconhecimento possuem resultados precisos para imagens de íris obtidas em ambientes ideais, onde há colaboração dos indivíduos, iluminação adequada, entre outras configurações que proporcionam na maioria dos casos uma localização e segmentação precisas da região alvo. Para realizar a tarefa de segmentação foram propostos diversos métodos que inicialmente baseavam-se nas fronteiras da íris com as regiões da pupila, esclera e pálpebras. Entre eles, os que mais se destacaram pelos bons resultados utilizam ou o operador íntegro diferencial [4] ou transformada de Hough [5]. O primeiro encontra os círculos que mais se adequam as fronteiras da íris com a esclera e a pupila [6]. O segundo faz uso de detecção de bordas de toda a imagem e, em seguida, encontra os parâmetros que mais adequa curvas as bordas da íris.

Entretanto, quando se trata de ambientes em condições não ideais de aquisição de imagens, os algoritmos existentes em geral não conseguem obter os mesmos resultados. Isso ocorre pois, em condições não ideais, surgem diversos desafios, como: oclusões causadas por cílios e pálpebras, baixa iluminação, reflexos de iluminação, ângulo de captura da câmera não ortogonal ao plano da íris, borramento devido a movimentação do indivíduo no momento da captura, uso de óculos, uso de lentes de contato, qualidade e ruído dos sensores de captura [6].

Devido a esses empecilhos no momento da captura das imagens que podem prejudicar a identificação das bordas da íris, os algoritmos baseados em *Deep Learning*, redes neurais profundas, vem sendo cada vez mais utilizados e já são considerados o estado-da-arte para muitas aplicações de segmentação, detecção e reconhecimento de imagens. Esses algoritmos se destacam pois são capazes de aprender e extrair quais características são mais importantes para definir se uma determinada região da imagem faz parte da íris ou não. Eliminando com isso a dependência de uma única característica, como por exemplo as bordas.

Quando o contexto de uso de redes neurais profundas envolve a manipulação de imagens é comum a utilização de Redes Neurais Convolutacionais (*Convolutional Neural Networks*, CNN). Essas redes baseiam-se na aprendizagem hierárquica aliada a técnicas de convolução que são comuns em procedimentos de processamento de imagens. À medida que as informações percorrem a rede, ocorre um aprendizado de estruturas cada vez mais abstratas que são de fundamental importância para estabelecer o que é mais importante dentro da tarefa proposta, seja ela detecção, segmentação

ou reconhecimento [7], [8].

Como as técnicas de *Deep Learning* que utilizam CNN realizam a aprendizagem de informações contidas em uma imagem, geralmente utiliza-se abordagens supervisionadas ou não supervisionadas. Nas abordagens não supervisionadas é fornecida para a rede neural uma imagem, porém não contendo a informação do que se trata a imagem ou das informações que ela contém. Dessa forma espera-se que a rede aprenda informações baseadas no contexto da imagem como um todo. Já na aprendizagem supervisionada é fornecida para uma rede neural não apenas uma imagem, mas o seu rótulo correspondente. Ou seja, caso a intenção seja o reconhecimento, a identificação do que se quer reconhecer. Se a tarefa for segmentação, é fornecida à rede qual região corresponde a área que deve ser segmentada [9].

Até o presente momento, alguns métodos utilizando CNN para segmentação de íris foram propostos e, em alguns casos, obtiveram melhores resultados do que as técnicas convencionais em ambientes de testes. Jalilian e Uhl [10] propuseram três variantes de uma técnica de *Deep Learning* que utiliza CNN conhecida como Autoencoder. Li et al [11] propuseram duas abordagens. A primeira utiliza blocos de camadas convolucionais e de pooling mapeando a saída desses últimos blocos para uma camada totalmente conectada. A segunda abordagem contém seis blocos semelhantes que têm cada uma de suas saídas combinadas em uma única camada que gera o resultado. Bazrafkan, Thavalengal e Corcoran [12] propõem a utilização de uma rede neural profunda semi paralela para gerar mapeamentos das regiões da íris quando são fornecidas imagens de baixa qualidade.

O presente trabalho tem como objetivo propor e implementar um Autoencoder convolucional, uma abordagem de *Deep Learning* utilizando CNN, para a segmentação da região da íris em imagens de olhos humanos. Além disso também foram realizados experimentos com o intuito de encontrar uma boa combinação de resultados que alcançasse ou superasse os resultados existentes na literatura. Algumas avaliações baseadas em métricas utilizadas em visão computacional e de segmentação de íris também foram realizadas.

II. MATERIAIS E MÉTODOS

A. Método Proposto

O método proposto utiliza uma abordagem denominada de Autoencoder ou encoder-decoder convolucional, que se baseia na codificação de informações de entradas em informações mais simples e mais relevantes seguidas de decodificações de tais informações com o intuito de construir um aproximação para a imagem alvo.

O Autoencoder proposto nesse trabalho é composto por 8 blocos principais, sendo 4 de codificação (*encoder*) da imagem de entrada e 4 de decodificação (*decoder*) para produção da imagem segmentada. Cada um dos dos 4 blocos de codificação são formados sequencialmente por uma camada convolucional com função de ativação ReLU (*Rectified Linear Unit*) seguida de uma camada *MaxPool* 2x2. Os blocos da etapa de decodificação são basicamente os mesmos da codificação, exceto pelo fato que as camadas de *MaxPool* são substituídas por filtros de deconvolução e aplicadas antes das camadas convolucionais correspondentes. Além disso, a última camada convolucional tem função de ativação linear. O otimizador ADAM (do inglês, *Adaptive Moment Estimation*) é utilizada para minimizar a função de custo MSE (*Mean Squared Error*), recomendada pela literatura para o problema de segmentação de imagens.

A ideia base do método proposto é minimizar o resultado da função de custo, dada uma imagem de íris, e outra imagem segmentada, fazendo assim com que o método aprenda as melhores características para identificar quais são os pixels correspondentes a região da íris e assim minimizar a diferença entre ambas.

B. Banco de Imagens e Aumento de Dados

No presente trabalho foram utilizados 3 bases de imagens de íris bem conhecidos na literatura. A primeira base é o



Fig. 1. Imagens pertencentes ao 3 bases de imagens. CASIA-Íris-Interval (esquerda), Ubiris v2 (centro) e IIT Delhi (direita).



Fig. 2. Imagens de gabarito correspondentes as imagens da figura 1.

CASIA-Íris-Interval (*casia4i*), que é composto de 2639 imagens e é um subconjunto da base de imagens CASIA v4.0 [13]. Também foi utilizado um subconjunto composto de 2250 imagens do Ubiris v2 [14]. A terceira base, IIT Delhi Íris Database version 1.0, foi utilizado em sua totalidade e é composto por 2240 imagens [15]. Ao todo, 7129 imagens foram utilizadas para treinamento e teste do método descrito neste trabalho (Fig. 1). Todas essas imagens foram manualmente segmentadas e disponibilizadas nos trabalhos de [16] e [17] (Fig. 2). Para o restante deste trabalho, essas imagens são denominadas imagens de gabarito.

Quando se utiliza algoritmos de aprendizagem é comum a divisão dos dados para poder avaliar durante o treinamento o quanto um algoritmo está aprendendo. Como as 3 bases de imagens têm uma quantidade de imagens próximas, optou-se por separar 60% das imagens de cada base para a etapa de treinamento (4277 imagens), 20% para etapa de validação (1426 imagens) e os 20% restante para a etapa de teste (1426 imagens). As imagens de validação são utilizadas durante o treinamento para verificar como o algoritmo desempenha a tarefa de segmentação, em imagens com as quais ele não foi treinado, e para ajustar hiperparâmetros. Ao final de todo o processo de treinamento é realizado uma nova verificação no conjunto de teste, que é de onde provêm os dados informados na seção de resultados.

Sabe-se que para atingir um bom desempenho em tarefas complexas, os algoritmos de *Deep Learning*, necessitam de grandes quantidades de dados. Para isso foram realizadas técnicas de aumento de dados que baseiam-se na aplicação de transformações espaciais nas imagens contidas no conjunto de treinamento e as mesmas transformações aos seus respectivos gabaritos. Neste trabalho, foram utilizadas as seguintes transformações: rotações entre ± 15 graus, deslocamento lateral de até 20% para esquerda ou para direita, deslocamento vertical de até 10% para cima ou para baixo, e espelhamento em torno do eixo vertical.

C. Métricas

Com o intuito de avaliar a performance do método proposto, foram utilizadas métricas comuns na literatura de segmentação de íris. Essas métricas abrangem medidas chave para avaliar tarefas de segmentação como: verdadeiros positivos (TP), falsos positivos (FP), falsos negativos (FN).

A primeira métrica é o coeficiente de similaridade de Jaccard (Equação 1), onde A corresponde ao conjunto de pixels segmentados pelo método como pertencentes a íris, e B corresponde ao conjunto de pixels pertencentes a íris na imagem de gabarito.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|} \quad (1)$$

Para entender a segunda métrica é preciso inicialmente definir outras duas métricas. A primeira é denominada Precisão (Equação 2) que informa quantos pixels considerados íris são realmente pertencentes a íris.

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (2)$$

A segunda métrica é o Revocação (Equação 3) que mede quantos dos pixels verdadeiramente da íris foram segmentados como íris pelo algoritmo.

$$\text{Revocação} = \frac{TP}{TP + FN} \quad (3)$$

Por intermédio de Precisão e Revocação pode-se definir a *f1-measure* (Equação 4), que é a média harmônica entre Precisão e Revocação.

$$f1 = \frac{2 \times \text{Precisão} \times \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (4)$$

A métrica *nice1* (Equação 5) é mais uma métrica avaliada e amplamente aceita para avaliação da precisão da segmentação da íris. A pontuação de erro de segmentação *nice1* calcula a proporção de pixels correspondentes que estão em discordância sobre toda a imagem da seguinte forma:

$$\text{nice1} = \frac{1}{c \times r} \sum_{c'} \sum_{r'} O(c', r') \otimes C(c', r') \quad (5)$$

Onde *c* e *r* são o número de colunas e linhas, respectivamente, das imagens de gabarito; $O(c', r')$ e $C(c', r')$ são respectivamente, pixels da imagem gerada pelo método e da imagem de gabarito; e \otimes é o operador lógico XOR.

A última métrica de erro de segmentação avaliada foi a *nice2* (Equação 6) e é composta pela média da soma do número de falsos positivos e falsos negativos.

$$\text{nice2} = \frac{FP + FN}{2} \quad (6)$$

Para coeficiente de similaridade de Jaccard e *f1-measure*, o melhor resultado é 1 (ou 100%) e o pior 0 (ou 0%). Já para os erros *nice1* e *nice2*, o melhor resultado é 0 e o pior 1.

Na implementação do método proposto neste trabalho foi utilizada a linguagem de programação Python e o framework Keras. Os testes executados foram realizados em uma máquina com 16GB (gigabyte) de memória RAM DDR4, processador AMD Ryzen™ 7 1700X (8C/16T, 20MB Cache, 3.4GHz a 3.8GHz com Max Turbo), placa de vídeo NVIDIA GeForce GTX 1060 6GB GDDR5 e sistema operacional Windows 10 64 Bits.

III. RESULTADOS E DISCUSSÃO

Essa seção expõe os resultados obtidos por intermédio da utilização do método proposto aplicado nas imagens pertencentes ao conjunto de teste. Como descrito na seção anterior, o método teve seu desempenho avaliado principalmente sobre 4 métricas: coeficiente de similaridade de Jaccard, *f1-measure*, *nice1* e *nice2*. A Tabela I apresenta os resultados da avaliação do método em relação as imagens do conjunto de teste, que é composto por 20% de imagens de cada uma das 3 bases de imagens citadas na seção anterior. O resultado visual do método pode ser observado na Fig. 3.

Também avaliou-se o resultado do método individualmente sobre cada base de imagens do conjunto de teste. Tais resultados podem ser vistos na Tabela II.

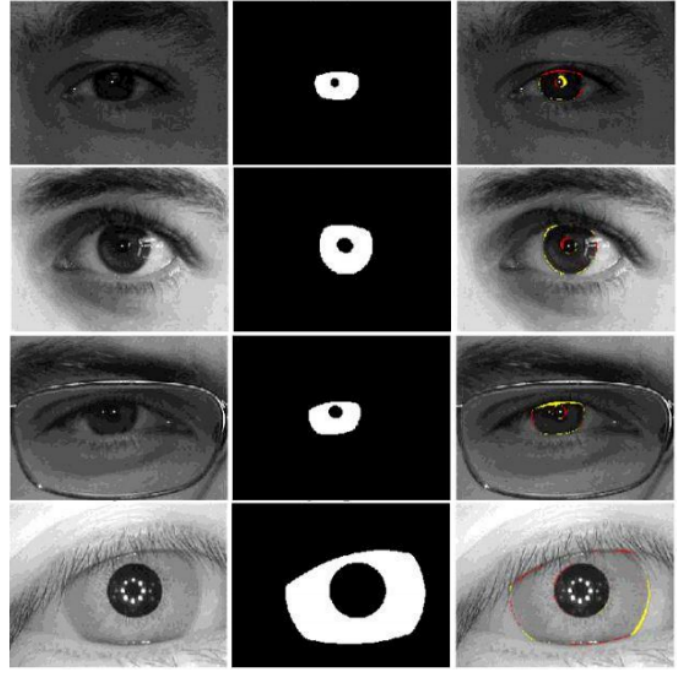


Fig. 3. Exemplos de resultados de segmentação. Imagens de entrada (esquerda), gabarito correspondente (centro). Na imagem da direita os pixels falsos positivos (vermelhos) e falsos negativos (amarelos), foram sobrepostos a imagem de entrada.

TABELA I
RESULTADOS DO MÉTODO PROPOSTO

Jaccard	f1-measure	nice1	nice2
93.47%	96.51%	0.0114	0.0202

De forma a potencializar o entendimento sobre os resultados, foram feitas comparações do resultado do método proposto com os resultados encontrados em artigos da literatura.

Ao compararmos o resultado da *f1-measure* sobre os resultados de imagens da base Ubiiris v2, pode-se identificar uma redução de mais de 59% do erro, sobre o melhor resultado encontrado na literatura. Quando a métrica de comparação é a *nice1*, o método proposto neste trabalho também foi superior aos métodos encontrados na literatura, reduzindo-a em pelo menos 12% o erro, em relação ao melhor resultado apresentado na literatura. Em relação a métrica *nice2*, houve uma redução de em mais de 73% do erro quando comparado ao melhor resultado (Tabela III).

Nas bases CASIA-Íris-Interval e IIT Delhi, foram feitas comparações em relação as métricas *nice1* e *nice2*, pois nenhum dos trabalhos comparados apresentou resultados referentes a métrica *f1-measure*. Para o CASIA-Íris-Interval, obteve-se uma redução do erro de aproximadamente 68% na *nice1* e 62% na *nice2* (Tabela IV). No base IIT Delhi a melhora foi ainda maior, reduzindo o erro em mais de 74% a *nice1* e 71% a *nice2* (Tabela V).

Vale salientar que é natural que todos os métodos apresentem um resultado inferior na base Ubiiris v2, pois essa base de imagens simula condições de imagens não cooperativas, contendo em alguns casos imagens onde a íris está totalmente oclusa, Fig. 4.

Outro ponto a ser destacado foi a dificuldade para obter bases de imagens de íris, tanto as imagens originais como os seus respectivos gabaritos, o que restringiu as etapas de treinamento, validação e teste, não permitindo comparações com outros trabalhos existentes na literatura.

TABELA II
RESULTADOS DO MÉTODO PROPOSTO POR BASE DE IMAGENS

	Jaccard	f1-measure	nice1	nice2
Ubiris v2	88.32%	93.56%	0.0079	0.0296
Casia v4	96.12%	98.02%	0.0125	0.0151
IIT Delhi	95.52%	97.70%	0.0137	0.0168

TABELA III
COMPARAÇÃO DO MÉTODO PROPOSTO COM MÉTODOS DA LITERATURA PARA A BASE UBIRIS V2

	[10]	[11]	[12]	Proposto
f1	84.07%	-	78.08%	93.56%
nice1	0.0306	0.0090	-	0.0079
nice2	0.1116	-	-	0.0296



Fig. 4. Imagem onde a região da íris encontra-se totalmente oculta (esquerda), gabarito correspondente a essa imagem (centro), e predição do método proposto (direita).

IV. CONCLUSÃO

O presente trabalho visou a implementação de um Autoencoder convolucional para segmentar a região da íris em olhos humanos e assim melhorar as etapas posteriores do processo de reconhecimento biométrico. É de suma importância que a segmentação seja realizada com um alto nível de qualidade, pois uma vez que erros aconteçam nessa etapa, eles serão propagados por todo o processo de tal maneira que não podem ser corrigidos ou recuperados.

Em ambientes de captura de imagens não-cooperativos, o desafio se torna ainda maior pela quantidade de empecilhos que podem ocorrer, em alguns casos dificultando até a própria capacidade humana de identificação de onde está localizada a região da íris em uma imagem. Diante disso, foram empregadas técnicas de *Deep Learning*, visto que os métodos propostos mais recentes vêm se utilizando de técnicas inovadoras nesse campo e com isso superando os métodos convencionais.

Como abordagem para solucionar o problema, foram utilizadas imagens tanto de ambientes cooperativos como não-cooperativos. Os resultados mostraram que o algoritmo teve um desempenho superior em comparação à alguns dos melhores resultados presentes na literatura em diferentes tipos de imagens e diferentes métricas. Tais resultados, comprovam a capacidade do método de extrair características que permitem identificar a íris.

Em trabalhos futuros, espera-se utilizar outras técnicas de *Deep Learning* para as demais etapas do processo de reconhecimento, como extração de características para identificação e o casamento de características. Outra alternativa seria a realização de um novo treinamento, dispondo de uma maior quantidade de imagens e em situações mais variadas.

AGRADECIMENTOS

Gostaríamos de agradecer em especial a Arnaldo Gualberto que auxiliou no presente trabalho, estimulando o estudo de técnicas de *Deep Learning* para a resolução do problema proposto.

REFERENCES

[1] J. Daugman, "Recognising persons by their iris patterns," in *Advances in biometric person authentication*. Springer, 2004, pp. 5–25.

TABELA IV
COMPARAÇÃO DO MÉTODO PROPOSTO COM MÉTODOS DA LITERATURA PARA A BASE CASIA V4

	[10]	[11]	[12]	Proposto
nice1	0.0391	-	-	0.0125
nice2	0.0407	-	-	0.0151

TABELA V
COMPARAÇÃO DO MÉTODO PROPOSTO COM MÉTODOS DA LITERATURA PARA A BASE IIT DELHI

	[10]	[11]	[12]	Proposto
nice1	0.0539	-	-	0.0137
nice2	0.0594	-	-	0.0168

[2] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Personal identification based on iris texture analysis," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 12, pp. 1519–1533, 2003.

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

[4] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE transactions on pattern analysis and machine intelligence*, vol. 15, no. 11, pp. 1148–1161, 1993.

[5] R. P. Wildes, "Iris recognition: an emerging biometric technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348–1363, 1997.

[6] K. W. Bowyer and M. J. Burge, *Handbook of iris recognition*. Springer, 2016.

[7] Y. LeCun, D. Touresky, G. Hinton, and T. Sejnowski, "A theoretical framework for back-propagation," in *Proceedings of the 1988 connectionist models summer school*, vol. 1. CMU, Pittsburgh, Pa: Morgan Kaufmann, 1988, pp. 21–28.

[8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[9] B. M. Wilamowski and J. Korniak, "Learning architectures with enhanced capabilities and easier training," in *Intelligent Engineering Systems (INES), 2015 IEEE 19th International Conference on*. IEEE, 2015, pp. 21–29.

[10] E. Jalilian and A. Uhl, "Iris segmentation using fully convolutional encoder-decoder networks," in *Deep Learning for Biometrics*. Springer, 2017, pp. 133–155.

[11] N. Liu, H. Li, M. Zhang, J. Liu, Z. Sun, and T. Tan, "Accurate iris segmentation in non-cooperative environments using fully convolutional networks," in *Biometrics (ICB), 2016 International Conference on*. IEEE, 2016, pp. 1–8.

[12] S. Bazrafkan, S. Thavalengal, and P. Corcoran, "An end to end deep neural network for iris segmentation in unconstrained scenarios," *arXiv preprint arXiv:1712.02877*, 2017.

[13] "Biometrics ideal test, casia.v4 database." [Online]. Available: <http://www.idealtest.org/dbDetailForUser.do?id=4>

[14] H. Proenca, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, "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*, vol. 32, no. 8, pp. 1529–1535, 2010.

[15] A. Kumar and A. Passi, "Comparison and combination of iris matchers for reliable personal authentication," *Pattern recognition*, vol. 43, no. 3, pp. 1016–1026, 2010.

[16] F. Alonso-Fernandez and J. Bigun, "Near-infrared and visible-light periocular recognition with gabor features using frequency-adaptive automatic eye detection," *IET Biometrics*, vol. 4, no. 2, pp. 74–89, 2015.

[17] H. Hofbauer, F. Alonso-Fernandez, P. Wild, J. Bigun, and A. Uhl, "A ground truth for iris segmentation," in *Pattern Recognition (ICPR), 2014 22nd International Conference on*. IEEE, 2014, pp. 527–532.