Person Authentication Based on the Difference of Deep Features Extracted from the Ocular and Face Regions

Marcelo Vilela Vizoni
Department of Computing - Faculty of Sciences
UNESP - São Paulo State University
Bauru, Brasil
vizonimarcelo@gmail.com

Aparecido Nilceu Marana
Department of Computing - Faculty of Sciences
UNESP - São Paulo State University
Bauru, Brasil
nilceu.marana@unesp.br

***Resumo*—This paper presents a new method for person authentication that relies on the fusion of two biometric authentication methods based, respectively, on ocular deep features and facial deep features. In our work, the deep features are extracted from the regions of interest by using a very deep CNN (Convolutional Neural Network). Another interesting aspect of our work is that, instead of using directly the deep features as input for the authentication methods, we use the difference between the probe and gallery deep features. So, our method adopts a pairwise strategy. Support Vector Machine classifiers are trained separately for each approach. The fusion of the ocular and the facial based methods are carried out in the score level. The proposed method was assessed with a facial database taken under uncontrolled environment and reached good results. Besides, the fusion strategy proposed in this work showed better results than the results obtained by each individual method.**

***Index Terms*—Person Authentication, Ocular Recognition, Face Recognition, Multibiometrics, Deep Learning, Convolutional Neural Networks.**

I. Introdução

Os avanços tecnológicos dos últimos anos acabaram criando um cenário onde a identificação automática de indivíduos se tornou uma grande necessidade. Os métodos tradicionais de identificação, como os baseados em posse (ex, cartões) ou conhecimento (ex, senhas), estão cada vez mais sendo deixados de lado para dar espaço à abordagem denominada Biometria [1], que busca a identificação dos indivíduos por meio de suas características biológicas ou comportamentais.

Uma das principais vantagens da utilização de sistemas biométricos é a maior segurança proporcionada contra fraudes, devido ao fato de que a identificação por meio de posse ou conhecimento pode ser facilmente burlada, pois as posses podem ser perdidas ou emprestadas e o conhecimento pode ser esquecido, inferido ou até mesmo compartilhado indevidamente pelo próprio indivíduo [2].

A partir deste contexto, diversos tipos de sistemas biométricos têm sido investigados e utilizados, abordando técnicas por reconhecimento de uma variedade de características biológicas ou comportamentais dos indivíduos. Após uma fase inicial de pesquisa e desenvolvimento de sistemas biométricos, na qual apenas dados obtidos em ambientes controlados eram possíveis de serem utilizados, o interesse passou a ser o desenvolvimento de sistemas capazes de operar também em ambientes não controlados, com aquisição à distância dos dados biométricos e sem a colaboração das pessoas sendo identificadas, o que aumentou significativamente os desafios do reconhecimento biométrico [3]. Nessas condições, a biometria facial é uma das que melhor atendem aos requisitos. A biometria periocular também vem ganhando espaço nesta área de pesquisa devido aos desempenhos aquém do desejável obtidos em alguns casos por sistemas biométricos baseados apenas nas características faciais [3]–[5].

Neste trabalho propomos um novo método para autenticação de pessoas que realiza a fusão de dois métodos de autenticação biométrica baseados, respectivamente, em características oculares profundas (extraídas das regiões dos olhos esquerdo e direito) e características faciais profundas (extraídas de toda a região da face). No método proposto, as características profundas são extraídas das regiões de interesse usando uma arquitetura de Rede Neural Convolucional (CNN, Convolutional Neural Network) muito profunda chamada VGG-Face [6]. Outro aspecto interessante do método proposto é que, ao invés de usar diretamente as características profundas como entrada para os métodos de autenticação, são utilizadas as diferenças entre as características *probe*, de busca, e as características *gallery*, armazenadas na base de dados. Ou seja, o método proposto é baseado nas diferenças existentes entre pares (*pairwise differences*) de características biométricas. Máquinas de vetores de suporte (SVM, Support Vector Machines) são treinadas separadamente em conjuntos de pares de características profundas oculares e faciais genuínas e impostoras a fim de aprender a reconhecer quando um par de características é genuino, ou seja, oriundo da mesma pessoa, ou impostor, ou seja, oriundo de pessoas distintas. A fusão dos métodos individuais (baseados nas características oculares e faciais) é realizada no nível de pontuação (usando a função soma).

As probabilidades fornecidas pelos classificadores SVM são usadas como pontuações que, portanto, variam entre 0 e 1.

II. Região Ocular

A região ocular refere-se a região do rosto que contém o olho e sua vizinhança, como os cílios, as pálpebras, as sobrancelhas e a pele ao seu redor. Características dessa região têm sido utilizadas para compor sistemas biométricos que buscam aspectos únicos em indivíduos, podendo ser utilizadas sozinhas ou em conjunto com a íris ou a face [3], uma vez que as informações da região ocular podem ser capturadas pelo mesmo sensor utilizado para capturar as imagens das íris e das faces. Desta forma, a Biometria da região ocular pode se apresentar como uma alternativa para situações em que haja dificuldades em reconhecer a íris ou a face, seja por obstrução da imagem ou pelo posicionamento inadequado do sensor no momento da captura. A Figura 1 mostra exemplos de imagens de regiões oculares.

Figura 1. Exemplos de imagens da região ocular [7]

III. Características Profundas

Devido à capacidade de reconhecer padrões complexos de forma holística e inspirada nos modelos biológicos, os métodos de aprendizado de máquina em profundidade têm ganhado popularidade. Segundo Deng e Dong [8], por extraírem informações de alto nível e autoaprendidas dos dados sendo tratados, as técnicas de aprendizado de máquina em profundidade conseguem obter alta abstração, grande poder de generalização e, consequentemente, grande robustez em suas aplicações.

Esses métodos têm sido aplicados com êxito em tarefas de reconhecimento de padrões em diversos domínios. As CNN (*Convolutional Neural Network*), por exemplo, têm demonstrado robustez às variações espaciais intraclasses, muito comuns nas aplicações da vida real, e têm sido bem sucedidas no reconhecimento de caracteres manuscritos [9], na detecção de objetos [10], na classificação de imagens em larga escala [11].

Convolutional Neural Networks (CNN) [9] são arquiteturas de aprendizagem profunda constituídas de camadas nas quais diferentes tipos de filtros (convolução e amostragem) são aplicados aos dados de entrada, inicialmente imagens bidimensionais. O resultado de uma dada camada serve como entrada para a camada acima até o topo da rede ser alcançado. Além da convolução e amostragem, camadas com neurônios totalmente conectados podem ser incluídas no topo da rede para classificação, geralmente realizando retificação de sinal ou aplicando a função de normalização softmax, neste caso, transformando seus valores de entrada em probabilidades de saída. Nas camadas superiores da rede são obtidas representações de alto nível da imagem original, mais robustas do que as informações de pixels brutos para muitas aplicações. A Figura 2 ilustra a arquitetura de uma Rede Neural Convolucional.

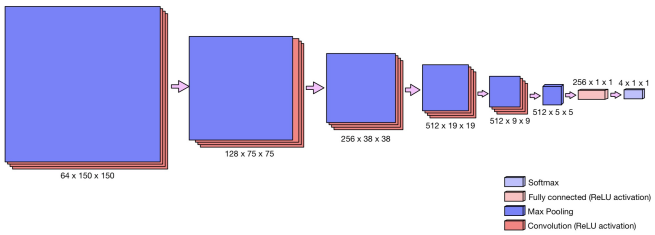

Figura 2. Arquitetura da Rede Neural Convolucional VGG16.

Este sucesso também tem sido verificado na área de Biometria [12]–[15]. No reconhecimento de faces, por exemplo, métodos baseados em CNN têm superado significativamente os métodos tradicionais baseados em características *handcrafted* [12], [16]. No reconhecimento de íris, por exemplo, a aplicação de características profundas (deep features), extraídas da VGG-Net [17], em um algoritmo de classificação simples, mostrou que as características aprendidas via treinamento para uma tarefa muito diferente de reconhecimento de objetos podem ser transferidas para o reconhecimento biométrico das íris, superando os melhores resultados obtidos por técnicas tradicionais [18].

Neste trabalho, utilizamos a rede VGGFace, uma rede neural convolucional pré-treinada com milhares de imagens de faces [6]. Como o objetivo é utilizar a rede para extração de características profundas, as informações que ela fornece são extraídas sem realizar a etapa de classificação, anteriormente à camada totalmente conectada, que fornece as características em um total de 512 núcleos de formato 7x7 [19].

IV. Método Proposto

O método proposto neste trabalho para a autenticação de pessoas utilizando aprendizado em profundidade tem duas etapas principais: a extração de características e a autenticação dos indivíduos. O método proposto é aplicado separadamente nas imagens das regiões oculares direita e esquerda, e também da região facial total. Posteriormente, é realizada a fusão.

A. Extração de Características

Esta etapa é responsável pela detecção, nas imagens de entrada, da região da face e das regiões oculares (esquerda e direita), e também pela extração das características profundas dessas regiões.

Como o objetivo deste trabalho é reconhecer indivíduos à distância em ambientes não controlados, é natural que as imagens capturadas apresentem, além da face do indivíduo, outras partes do corpo. Assim, é primordial que se utilize um método robusto para a detecção da região facial e também das regiões oculares esquerda e direita nas imagens capturadas.

Para a detecção das regiões oculares e faciais foi utilizado neste trabalho o método MTCNN (*Multitask Cascaded Convolutional Networks*), que apresenta resultados muito bons quando comparado a outros métodos propostos na literatura [20]. O método MTCNN, além de possibilitar a segmentação do rosto realizando o enquadramento da face, também retorna as informações de posicionamento de olhos, nariz e boca na imagem. Essas informações possibilitam que as regiões oculares do olho esquerdo e direito sejam facilmente detectadas utilizando heurísticas baseadas na antropometria da face humana. A Figura 3 ilustra o processo da aplicação do MTCNN em uma imagem da base de dados FRGC [21] com a segmentação da face, e posteriormente a determinação das regiões oculares da esquerda e direita.

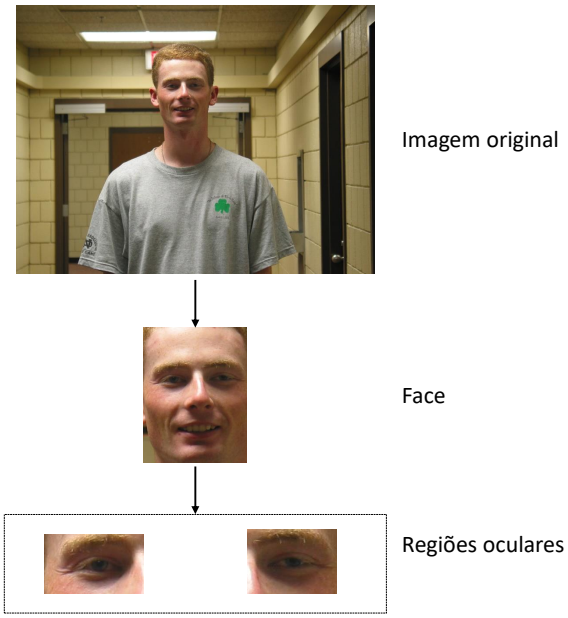

Figura 3. Aplicação do método MTCNN [20] para a segmentação da região facial e das regiões oculares dos olhos esquerdo e direito em uma imagem da base de dados FRGC (*Face Recognition Grand Challenge*) [21].

Como já foi mencionado, neste trabalho utilizamos a rede VGGFace para a extração de características profundas. Essas características foram obtidas da VGGFace, antes da etapa de classificação, no formato de 512x7x7. Portanto, os 512 núcleos formam a representação das características profundas obtidas das regiões oculares e faciais.

B. Autenticação dos Indivíduos

Esta etapa é responsável pela autenticação biométrica do indivíduo a partir dos vetores de características obtidos na etapa de extração de características. A autenticação é realizada no modo de autenticação, no qual o objetivo principal é verificar se as características biométricas obtidas do indivíduo no momento da autenticação, a partir da imagem *probe*, coincidem com as características biométricas daquele indivíduo armazenadas no banco de dados, obtidas da imagem *gallery*.

A maioria dos trabalhos que realizam autenticação extrai características independentemente de cada imagem. Nestes casos, as relações existentes entre os vetores de características não são modeladas a priori para os estágios de treinamento e de classificação. Assumindo a hipótese de que a modelagem das relações existentes entre vetores de características pode ser útil para aumentar a robustez e o desempenho das tarefas de autenticação, neste trabalho é proposta uma abordagem *pairwise*, baseada nas relações existentes entre pares de vetores de características. Para tanto, utiliza-se o vetor diferença, obtido da subtração das características *probe* e *gallery* dos indivíduos sendo autenticados. Sendo assim, o vetor é calculado pela diferença entre cada um dos 512 núcleos de 7x7, formando um vetor diferença de 512 características para cada comparação.

O referido vetor diferença é apresentado ao classificador no momento da autenticação para que este decida se a comparação em curso é genuína (os vetores *probe* e *gallery* são do mesmo indivíduo) ou impostora (os vetores *probe* e *gallery* são de indivíduos distintos).

Os classificadores utilizados no método proposto neste trabalho são máquinas de vetores de suporte (SVM, *Support Vector Machine*) [22]. Como estes classificadores fornecem as probabilidades dos vetores de características pertencerem às classes, os valores das probabilidades são utilizados como sendo as pontuações das comparações, permitindo, desse modo, o cálculo das curvas ROC (Receiver Operating Characteristics) e dos valores de EER (Equal Error Rate).

A Figura 4 apresenta um diagrama do método proposto para realizar a autenticação biometrica de pessoas baseada na estratégia *pairwise*. Nesta diagrama está ilustrado o caso específico da autenticação baseada na região ocular do olho esquerdo. O mesmo processo é realizado para a região ocular do olho direito e para a região que engloba toda a face.

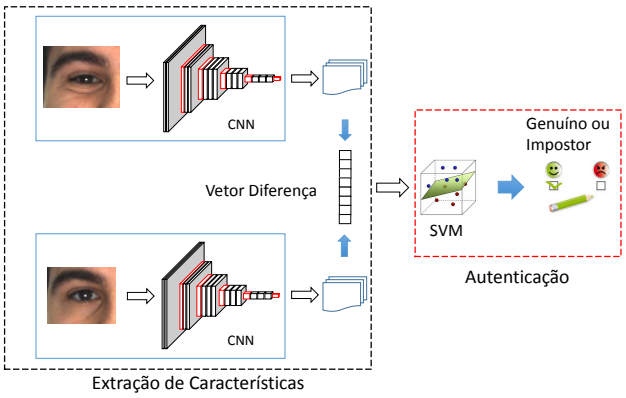

Figura 4. Diagrama do método proposto para autenticação biométrica de pessoas baseada na estratégia *pairwise* e nas características profundas obtidas, neste exemplo, a partir da região ocular do olho esquerdo da pessoa.

C. Fusão

Um sistema biométrico baseado em duas ou mais características biométricas é chamado de sistema biométrico multimodal [23]. Seu uso pode ser uma boa alternativa para melhorar o reconhecimento de pessoas, em cenários em que os dados de entradas podem sofrer variações ou ruídos.

Figura 5. Diferentes imagens do conjunto de dados FRGC (*Face Recognition Grand Challenge*) [21] capturadas de um mesmo indivíduo, à esquerda imagens capturadas em ambiente controlado, e a direita imagens capturas em ambientes não controlados.

Segundo Ross e Jain [24], a etapa de fusão de dois tipos de biometrias distintas pode ocorrer em três níveis distintos do processo de reconhecimento de um sistema biométrico:

- **Nível de características:** As informações obtidas por meio dos dados capturados por diferentes sensores são transformadas em vetores de características, e seus descritores combinados, como por exemplo por meio de uma concatenação. Um novo vetor de características com mais dimensões é obtido para representar o indivíduo, ou também técnicas de redução de dimensionalidade podem ser aplicadas;
- **Nível de pontuação:** Cada sistema oferece uma pontuação de similaridade entre os descritores analisados, e essas pontuações são combinadas para afirmar ou não a veracidade do modelo sendo analisado em questão;
- **Nível de decisão:** Baseado na pontuação de similaridade própria, cada sistema aceita ou rejeita a identidaade da pessoa. Após, um esquema de decisão (por exemplo, maioria de votos) é utilizado para se chegar a uma decisão única final.

Neste trabalho foi adotada a fusão no nível da pontuação, utilizando-se a função soma, visando aproveitar as pontuações (probabilidades retornadas pelos três classificadores SVM) obtidas para cada região de interesse: região ocular do olho esquerdo, região ocular do olho direito e região de toda a face.

V. Resultados Experimentais

O conjunto de dados FRGC (Face Recognition Grand Challenge) [21] foi criado na Universidade de Notre Dame e consiste em milhares de imagens coloridas de alta resolução (1200×1400 pixels), capturadas em ambientes controlados e não controlados. O subconjunto controlado foi capturado em um estúdio sob iluminação uniforme, onde os indivíduos eram obrigados a ficar parados enquanto olhavam diretamente para a câmera e ensaiavam diferentes expressões. Quanto ao conjunto não controlado, as imagens foram obtidas em diferentes cenários, desconsiderando o fundo e a iluminação. A Figura 5 mostra um exemplo de diferentes imagens capturadas de um mesmo indívíduo, à esquerda imagens capturadas em ambiente controlado, e a direita imagens capturas em ambientes não controlados..

O conjunto de dados FRGC consiste em seis diferentes experimentos, com comparações variadas entre indivíduos, em imagens 2D e 3D. Em nosso experimento abordamos o experimento 4, que consiste em testar a comparação entre uma única imagem de um indivíduo em ambiente controlado, com uma imagem em ambiente não controlado. Todos os 370 indivíduos do subconjunto de imagens 2D "fall 2013"foram selecionados aleatoriamente para realização do experimento. 12 imagens de cada indivíduo foram selecionadas aleatoriamente para gerar os vetores diferenças que treinaram o classificador SVM.

As imagens do conjunto de dados FRGC foram submetidas ao método de detecção de face MTCNN. Para cada imagem foram detectados o enquadramento da face, e o posicionamento dos olhos que são de interesse para realizar a segmentação das regiões oculares. Com estas informações, foram realizados recortes nas imagens originais da base de dados para se obter imagens contendo apenas as regiões faciais, e, posteriormente, com base nas distâncias entre os olhos, foram realizados novos recortes para se obter imagens contendo apenas as regiões oculares. Conhecendo-se as posições centrais de cada olho e os limites de cada face, é possível determinar (e recortar) a região de interesse ao redor de cada olho.

As imagens contendo apenas as regiões faciais e as imagens contendo apenas as regiões oculares foram submetidas a VGG-Face para a obtenção das características profundas (*deep features*) e a subsequente obtenção dos vetores de características.

Os conjuntos de vetores diferença das classes genuino e impostor foram obtidos, respectivamente, por meio da subtração de pares de vetores de características profundas oriundos

de diferentes imagens do mesmo indivíduo, e por meio da subtração de pares de vetores de características profundas oriundos de diferentes indivíduos. A etapa de teste, seguindo a protocolo do experimento 4 da base de dados FRGC, uma amostra em ambiente controlado e uma amostra em ambiente não controlado foram selecionadas para gerar os vetores diferenças. E da mesma forma que na etapa de treinamento, comparações intra-classe geraram descritores genuínos, e comparações inter-classe descritores impostores.

O treinamento supervisionado dos classificadores SVM foi promovido por meio da submissão dos vetores (amostras) das classes genuino e impostor do conjunto de treinamento, indicando-se os rótulos das classes de origem das amostras.

Após o treinamento, os classificadores foram utilizados na etapa de teste para classificar os vetores (amostras) das classes genuino e impostor do conjunto de teste. A fase de teste se deu basicamente pela apresentação de cada uma das amostras do conjunto de teste aos classificadores, que ficaram especializados em fornecer as probabilidades (pontuações) de cada amostra pertencer à classe genuíno.

Tendo a saída do classificador (probabilidade) para cada amostra é possível calcular as curvas ROC (Receiver Operating Characteristics) e as taxas de erro igual EER (Equal Error Rate).

A Figura 6 mostra a curva ROC obtida com a aplicação no método proposto nas imagens da base de dados FRGC. Para cada característica biométrica e cada abordagem (individual ou fundida) é apresentado o valor do EER (*Equal Error Rate*), que mostra a pontuação para a qual os valores de FAR (False Accept Rate) e FRR (False Reject Rate) se igualam, sendo que quanto menor é o valor de EER melhor é o desempenho do sistema biométrico. A Tabela I mostra os valores de EER obtidos.

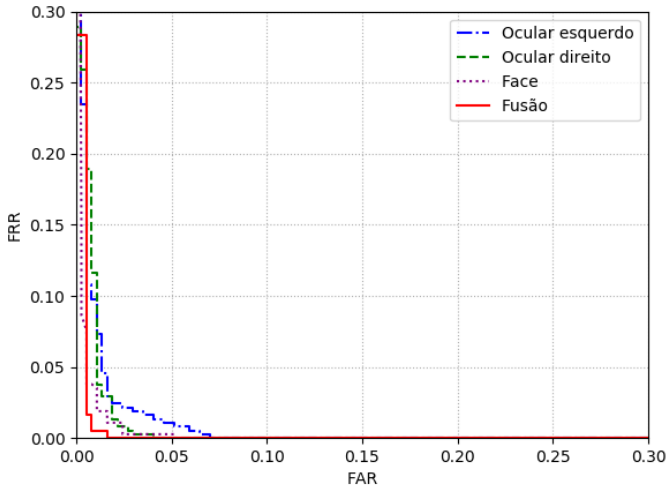

Figura 6. Curvas ROC obtidas com os experimentos realizados.

Entre as abordagens utilizando somente a região ocular a que obteve melhor resultado foi da região ocular direita com EER = 1,89%, e no geral, a abordagem utilizando a fusão entre toda a face com as regiões oculares com EER = 0,81%, superando o resultado de EER = 1,62% obtido pela abordagem utilizando somente a face. Os resultados obtidos mostram um bom desempenho do método proposto, levando em consideração a abordagem do experimento 4 da base de dados FRGC, que analisa imagens capturadas em ambientes controlados com imagens em ambientes não controlados.

Tabela I
RESULTADOS CONJUNTO DE DADOS FRGC

Abordagem	EER
Ocular direito	1,89%
Ocular esquerdo	2,25%
Face	1,62%
Fusão	0,81%

É possível encontrar outros trabalhos na literatura que utilizaram o experimento 4 da base de dados FRGC para gerar resultados em biometria facial. Em [25] os autores propõem um método para extração de características baseado em *Gabor Wavelets*, e um classificador SVM para fazer a autenticação facial, atingindo no melhor caso um resultado de 13% de EER. Em [26] é proposto uma nova estratégia como descritor facial chamado *Dual-Cross Patterns*, buscando trazer robustez a cenários que hajam variações de expressões faciais e iluminação, atingindo no melhor caso um resultado de 93,39% de taxa de verificação

VI. CONCLUSÕES

Pelos resultados obtidos é possível observar que o novo método proposto para autenticação de pessoas baseado na estratégia de *pairwise* de características profundas mostrou-se bastante efetivo para as três características biométricas alvo deste estudo: região ocular do olho direito, região ocular do olho esquerdo e região facial total, levando em consideração a base dados desafiadora utilizada. Os resultados também mostraram que as características das regiões oculares, ainda pouco exploradas nos sistemas biométricos, mesmo com taxas de erro mais elevadas que a face, podem ser úteis para realizar a autenticação biométrica em situações que toda a face não pode ser capturada. Observa-se, por fim, que a fusão dos métodos baseados nas características profundas obtidas das regiões oculares com o método baseado nas características profundas obtidas da região facial melhorou ainda mais os resultados obtidos individualmente pelos métodos.

Com isso, conclui-se que as características das regiões oculares, que ainda são pouco exploradas como identificadores biométricos, podem ser de grande valia na composição de sistemas multibiométricos mais robustos para a autenticação de pessoas à distância, em ambientes pouco controlados, e nos quais há pouca ou nenhuma colaboração dos indivíduos com o processo de autenticação.

VII. AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil

(CAPES) - Código de Financiamento 001. Os autores também agradecem à NVIDIA Corporation (GPU Grant Program) pela Titan xp concedida.

REFERÊNCIAS

[1] A. K. Jain, A. A. Ross, and K. Nandakumar, "Introduction," in *Introduction to Biometrics*, pp. 1–49, Springer, 2011.

[2] S. Pankanti, R. M. Bolle, and A. Jain, "Biometrics: The future of identification [guest eeditors' introduction]," *Computer*, vol. 33, no. 2, pp. 46–49, 2000.

[3] G. Santos and H. Proença, "Periocular biometrics: An emerging technology for unconstrained scenarios," in *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2013 IEEE Workshop on*, pp. 14–21, IEEE, 2013.

[4] F. Alonso-Fernandez and J. Bigun, "Periocular biometrics: databases, algorithms and directions," in *2016 4th International Conference on Biometrics and Forensics (IWBF)*, pp. 1–6, March 2016.

[5] I. Nigam, M. Vatsa, and R. Singh, "Ocular biometrics: A survey of modalities and fusion approaches," *Information Fusion*, vol. 26, pp. 1–35, 2015.

[6] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *British Machine Vision Conference*, 2015.

[7] C. N. Padole and H. Proenca, "Periocular recognition: Analysis of performance degradation factors," in *2012 5th IAPR International Conference on Biometrics (ICB)*, pp. 439–445, March 2012.

[8] L. Deng and Y. Dong, "Deep learning: methods and applications," *Foundations and Trends in Signal Processing*, vol. 7, no. 3–4, pp. 197–387, 2014.

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

[10] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '14, (Washington, DC, USA), pp. 580–587, IEEE Computer Society, 2014.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, June 2015.

[12] Y. Sun, X. Wang, and X. Tang, "Deep learning face representation from predicting 10,000 classes," in *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '14, (Washington, DC, USA), pp. 1891–1898, IEEE Computer Society, 2014.

[13] W. Wang, J. Yang, J. Xiao, S. Li, and D. Zhou, "Face recognition based on deep learning," in *International Conference on Human Centered Computing*, pp. 812–820, Springer, 2014.

[14] A. Meraoumia, F. Kadri, H. Bendjenna, S. Chitroub, and A. Bouridane, "Improving biometric identification performance using pcanet deep learning and multispectral palmprint," in *Biometric Security and Privacy*, pp. 51–69, Springer, 2017.

[15] X. Sun, P. Wu, and S. C. Hoi, "Face detection using deep learning: An improved faster rcnn approach," *arXiv preprint arXiv:1701.08289*, 2017.

[16] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1701–1708, June 2014.

[17] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[18] S. Minaee, A. Abdolrashidiy, and Y. Wang, "An experimental study of deep convolutional features for iris recognition," in *2016 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pp. 1–6, Dec 2016.

[19] N. Ballas, L. Yao, C. Pal, and A. Courville, "Delving deeper into convolutional networks for learning video representations," *arXiv preprint arXiv:1511.06432*, 2015.

[20] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[21] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 1, pp. 947–954 vol. 1, June 2005.

[22] N. Guenther, M. Schonlau, *et al.*, "Support vector machines," *Stata J*, vol. 16, no. 4, pp. 917–937, 2016.

[23] T. Panchal and A. Singh, "Multimodal biometric system," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 5, 2013.

[24] A. Ross and A. Jain, "Information fusion in biometrics," *Pattern recognition letters*, vol. 24, no. 13, pp. 2115–2125, 2003.

[25] B. Chinmay, B. Rupesh, B. Nikhil, and R. Milind, "Face identification." IJESC 2017, Research article volume 7 , No 5, 2017.

[26] C. Ding, J. Choi, D. Tao, and L. S. Davis, "Multi-directional multi-level dual-cross patterns for robust face recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 3, pp. 518–531, 2015.