

# PREVENÇÃO DE FALSIFICAÇÃO EM SISTEMAS DE RECONHECIMENTO FACIAL: UMA PROPOSTA BASEADA EM CLUSTERIZAÇÃO

Caio César Rodrigues Garcez, Gilmar dos Santos Marques, Edna Dias Canedo, Bruno Garcia Praciano, Francisco Lopes de Caldas Filho e Fábio Lúcio Lopes de Mendonça  
*Pós-graduação Profissional em Engenharia Elétrica – PPEE – Departamento de Engenharia Elétrica, Faculdade de Tecnologia, Universidade de Brasília (UnB), Brasília, Brasil, Zip Code 70910-900*

## RESUMO

Embora o desempenho de Sistemas de Reconhecimento Facial (SRF) na identificação de usuários seja atualmente satisfatório, no que se refere às aplicações de controle de acesso, o aumento do uso de biometria falsa representa uma ameaça significativa à confiabilidade e segurança desses sistemas biométricos. Garantir a autenticidade dos usuários e desenvolver técnicas robustas de detecção de autenticidade, do inglês *liveness detection*, são aspectos cruciais para aprimorar a resistência contra tentativas de falsificação, do inglês *spoofing*, garantindo, assim, a eficácia contínua desses sistemas. O desafio de diferenciar uma biometria autêntica de uma biométrica falsa tem se tornado objeto de pesquisa e desenvolvimento. Este artigo propõe como mecanismo de prevenção de falsificação em sistemas de reconhecimento facial a aplicação de técnicas de incorporação, do inglês *embeddings*, seguidas de formas de agrupamento conforme o algoritmo t-distributed Stochastic Neighbor Embedding (t-SNE) para formação clusters a partir dessas incorporações. Ao contrário de outras técnicas de redução de dimensionalidade, o t-SNE mantém a estrutura local e global dos dados, tornando-o ideal para interpretar a distribuição complexa de imagens faciais. Nossa abordagem vai além da simples identificação de faces, estendendo-se à compreensão de toda a imagem, capturando assim as sutilezas e discrepâncias ambientais potencialmente associadas a tentativas de falsificação. Nossos resultados experimentais revelam que a combinação de incorporações de imagens e clusterização com t-SNE melhora significativamente a detecção de ataques de falsificação em sistemas de reconhecimento facial, representando um passo substancial no combate às tentativas de falsificação.

## PALAVRAS-CHAVE

Sistemas de Reconhecimento Facial, Detecção de Autenticidade, Técnicas de Clusterização

## 1. INTRODUÇÃO

A utilização de Sistemas de Reconhecimento Facial (SRF) tem sobressaído em relação a outros sistemas biométricos, como o reconhecimento de voz e a identificação por impressões digitais, principalmente devido às suas vantagens, que incluem a ausência de contato e o fato de não ser intrusivo. A tecnologia de Reconhecimento Facial (RF) tem apresentado progresso significativo nos últimos anos, sendo amplamente utilizada como uma etapa adicional nos processos de identificação em aplicações voltadas à segurança e vigilância.

A biometria baseada em SRF se tornou um foco significativo de pesquisa e desenvolvimento. Neste cenário, a ocorrência de tentativas de falsificação, do inglês *spoofing*, nas quais um indivíduo se disfarça como outra pessoa para obter acesso a um sistema, apresentam um desafio considerável à integridade desses sistemas e aos processos de autenticação (Dawson et al., 2015).

Nesse contexto, o presente trabalho propõe um mecanismo de prevenção a falsificações em SRF, utilizando técnicas de incorporação, do inglês *embeddings*, em conjunto com a estratégia de agrupamento conforme o algoritmo t-distributed Stochastic Neighbor Embedding (t-SNE) (Van Der Maaten et al., 2008) para formar clusters a partir dessas incorporações. Distintamente de outras técnicas de redução de dimensionalidade, o t-SNE preserva a estrutura local e global dos dados, tornando-se ideal para decifrar a distribuição complexa de imagens faciais. A abordagem apresentada vai além da simples identificação de faces, ampliando-se para a análise integral da imagem, capturando desta forma as nuances e as discrepâncias ambientais potencialmente

associadas às tentativas de falsificação. Os resultados experimentais apontam que a combinação das incorporações de imagens com a clusterização via t-SNE otimiza significativamente a detecção de ataques de falsificação em sistemas de reconhecimento facial. Este avanço representa um progresso substancial no combate a práticas de falsificação.

Este artigo é estruturado em cinco seções principais. A primeira delas, a Introdução, oferece uma contextualização sucinta da contribuição proposta por este trabalho. A seção subsequente, o Referencial Teórico e Trabalhos Correlatos, introduz as técnicas de reconhecimento facial baseadas em incorporações e técnicas de clusterização. Na terceira parte, discutimos a proposta deste trabalho e os passos adotados para a obtenção da quarta seção, a qual aborda os Resultados. Essa seção destaca as vantagens do uso da nossa proposta em comparação com técnicas convencionais de reconhecimento facial. Finalmente, o artigo se encerra com a seção de Conclusão, onde é feito um resumo do estudo, apresentando as conclusões obtidas a partir dos resultados, e se explora possíveis caminhos para pesquisas futuras.

## 2. REFERENCIAL TEÓRICO E TRABALHOS RELACIONADOS

### 2.1 Sistemas de Reconhecimento Facial (SRF)

Métodos de Reconhecimento Facial (RF) baseados em aprendizado de máquina englobam uma diversidade de técnicas voltadas à identificação, classificação e reconhecimento de rostos em imagens. Abordagens comumente empregadas na literatura incluem a Análise de Componentes Principais (PCA, do inglês Principle Component Analysis – ABDULRAHMAN) (Muzammil et al., 2015), Máquinas de Vetores de Suporte (SVM, do inglês Support Vector Machines) (Noble, 2006), Análise Discriminante Linear (LDA, do inglês Linear Discriminant Analysis) (Xanthopoulos et al., 2013), Redes Neurais Convolucionais (CNNs, do inglês Convolutional Neural Networks) (O'Shea et al., 2015) e Redes Neurais Convolucionais Multitarefa (MTCNN, do inglês Multi-task Cascaded Convolutional Networks) (Li et al., 2020). Para uma revisão literária detalhada de cada abordagem, a referência (Aria et al., 2020) é recomendada ao leitor. Esta seção irá tratar de um processo compartilhado por essas técnicas: a formação de incorporações, do inglês *embeddings*, que são representações condensadas de imagens de rostos em um espaço de dimensões reduzidas.

No contexto de RF, *embeddings* são vetores de alta dimensão que carregam informações de uma imagem de rosto. Estes elementos são obtidos através de um processo usualmente realizado por CNNs em que os dados de uma imagem são transformados em uma representação numérica que mantém a essência dos traços faciais. Este processo inicia com a aplicação de várias camadas de convolução na imagem de entrada em que são extraídas características de baixo nível, como bordas e texturas, para em seguida progredir para características de alto nível, como formas de olhos, nariz e boca (Kazemi et al., 2014).



Figura 1. Processo de obtenção do vetor associado ao rosto de um usuário. Figura adaptada de (Hoàng et al., 2023)

Cada imagem de rosto é então codificada em um vetor (veja a Figura (1)), resultando no *embedding*. Esta representação serve como uma assinatura única de um rosto. A dimensão do vetor é reduzida quando comparada com a representação original da imagem, facilitando o processamento computacional e o armazenamento de dados. Além disso, a geometria do espaço de *embeddings* considera que rostos semelhantes resultam em vetores próximos uns dos outros, enquanto rostos diferentes resultam em vetores distantes. Isso fornece alta eficiência no processo de comparação entre rostos, facilitando a identificação e classificação em sistemas de RF.

## 2.2 Mecanismos de Clusterização

A clusterização, também conhecida como agrupamento, é um método de aprendizado de máquina não supervisionado que possibilita a segmentação de um conjunto de objetos em grupos, do inglês *clusters*. Esta abordagem visa organizar elementos com base em critérios de semelhança preestabelecidos, de tal maneira que objetos com alta similaridade sejam acomodados no mesmo cluster, enquanto objetos menos similares sejam distribuídos em clusters diferentes (Ochi et al., 2004). No contexto de RF, este conceito é aplicado ao agrupamento de *embeddings*, o que possibilita a coleção de rostos similares e consequentemente, permite a análise e a comparação entre rostos.

Um dos algoritmos comumente usados para visualizar e compreender a estrutura de clusters em dados de alta dimensão é o t-SNE (*t-Distributed Stochastic Neighbor Embedding*) (Van Der Maaten et al., 2008). O t-SNE preserva tanto a estrutura local (a proximidade de pontos de dados semelhantes) quanto a global (a relação entre grupos de pontos de dados), tornando-o ideal para entender a distribuição complexa de *embeddings* de imagens faciais. Isso permite que pesquisadores e engenheiros visualizem a eficácia de seus algoritmos de reconhecimento facial e otimizem seu desempenho.

## 2.3 Trabalhos Relacionados

Diversos métodos têm sido propostos na literatura para a prevenção de falsificação em sistemas de reconhecimento facial. Nesta seção, daremos destaque a algumas abordagens notáveis que formam o atual estado da arte nesta área de pesquisa.

(Erdogmus et al., 2013) exploraram em detalhes a vulnerabilidade de sistemas de reconhecimento facial 2D a falsificações com máscaras faciais 3D específicas do sujeito. Além disso, esse estudo emprega contramedidas baseadas em Padrões Binários Locais, do inglês *Local Binary Patterns*, utilizando dados de cor e profundidade. Para tal propósito, foi introduzida a base de dados 3DMAD, do inglês *3D Mask Attack Database*, a primeira base de dados 3D de *spoofing* disponível ao público. Neste contexto, o trabalho se concentra especificamente em tentativas de falsificação baseadas no uso de máscaras. O trabalho proposto neste artigo leva em conta diversas tentativas de falsificação apresentadas na base de dados (Zhang et al., 2020).

(Parkin et al., 2019) introduziram uma proposta inovadora de arquitetura de rede neural para a prevenção de falsificação. Esta abordagem é fundamentada na estrutura de dados de imagens e promove a agregação de características em vários níveis de camadas de rede. No entanto, a metodologia Parkin não contempla a utilização de clusterização como etapa adicional no processo decisório relacionado à autenticidade das imagens. Neste artigo, reforçamos a relevância deste procedimento na autenticação de imagens, posicionando-o como um elemento essencial em nossa abordagem.

(Gang et al., 2007) propuseram uma técnica em tempo real para detectar a autenticidade durante o processo de reconhecimento facial, concentrando-se na identificação de piscadelas espontâneas. Esta abordagem, embora inovadora e não intrusiva, apresenta uma dependência significativa em gravações de vídeo e múltiplas capturas, em contraponto ao uso de fotografias estáticas. Este é um ponto de restrição que o presente trabalho visa abordar e superar, apresentando soluções que operam de maneira eficaz com um único quadro de imagem estática.

Dentro do extenso repertório de estratégias já existentes na literatura para combater a falsificação em reconhecimento facial, nosso estudo é voltado a aspectos ainda não abordados, principalmente no cenário de controle de acesso. Apresentamos uma metodologia que não só é mais generalista, mas também versátil, almejando elevar a robustez e a adaptabilidade dos Sistemas de Reconhecimento Facial, tornando-os mais eficazes e aplicáveis em ambientes de controle de acesso.

## 3. MÉTODO PROPOSTO

Neste trabalho, apresentamos um método voltado para a prevenção de falsificação em sistemas de reconhecimento facial, apoiado em um conjunto de dados conhecido como CelebA-Spoof (Zhang et al., 2020). Esta base de dados é reconhecida por sua extensa coleção de imagens e anotações detalhadas, específicas para

a problemática da prevenção de falsificação facial. Na Figura (2), apresentamos um exemplo de uma amostra extraída deste conjunto de dados, que utilizamos como base para desenvolver e testar a proposta deste trabalho.



Figura 2. A imagem à esquerda representa uma face autêntica da base de dados CelebA-Spoof, enquanto a imagem à direita exemplifica uma tentativa de falsificação

A abordagem proposta nesse trabalho inicia com a obtenção de um *embedding* da imagem de entrada do usuário. Para esse fim, recorremos a modelos de transformadores, como o CLIP da OpenAI (Radford et al., 2021) que tem demonstrado um desempenho notável na geração de representações vetoriais a partir de imagens. Além do CLIP, existem outras ferramentas eficientes no contexto da geração de *embeddings*, tais como o DeepFace (Taigman et al., 2014) e o FaceNet (Schroff et al., 2015).

Após a obtenção dos vetores, um banco de dados especificamente projetado para lidar com dados vetoriais é empregado. A tecnologia escolhida foi o Pinecone DB (Sahoo et al., 2023). Esta base de dados possui desempenho eficiente na gestão de dados vetoriais em larga escala, além de disponibilizar recursos para pesquisa e comparação de vetores. Este banco de dados foi escolhido em razão da sua capacidade de escalabilidade, um atributo essencial caso o objetivo seja o desenvolvimento de uma aplicação robusta e escalável a partir do método proposto neste trabalho.

Posteriormente, os vetores armazenados no banco de dados são recuperados para realizar a clusterização baseada no algoritmo t-SNE. Esse método é capaz de agrupar os vetores de maneira eficiente, preservando a estrutura dos dados e facilitando a análise das semelhanças e diferenças entre as imagens de rostos. Os clusters resultantes são associados às imagens autênticas e às falsificadas, de forma a facilitar a detecção de tentativas de *spoofing*.

Na etapa final, ao considerar a verificação de uma nova imagem, o algoritmo *k-means* (Pham et al., 2005) é empregado para determinar o cluster ao qual a imagem em avaliação mais se assemelha. Considerando a associação prévia de clusters, o emprego do algoritmo permite decidir se a nova imagem representa uma tentativa de falsificação. A Figura (3) apresenta a arquitetura do método proposto.

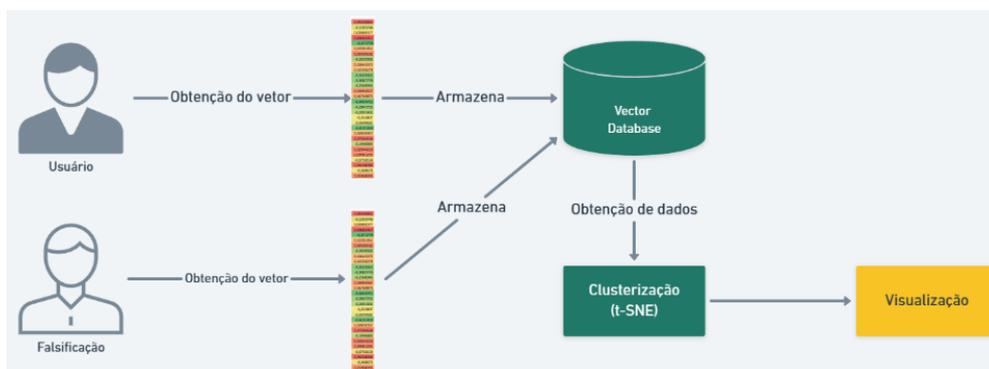


Figura 3. Arquitetura do método proposto

## 4. RESULTADOS

### 4.1 Métodos Tradicionais de Reconhecimento Facial

Na presente seção, buscamos ilustrar o desempenho de nosso método de reconhecimento facial em relação às técnicas tradicionais, utilizando uma subamostra do conjunto de dados da base CelebA-Spoof (Zhang et al., 2020). Vale ressaltar que os resultados aqui apresentados focam especificamente na correta clusterização, um elemento chave para a classificação precisa de rostos, dentro deste subconjunto do dataset.

Na primeira parte do procedimento experimental, a abordagem tradicional de reconhecimento facial é simulada para posteriormente servir de comparação com o método proposto. Esta abordagem envolve a geração de *embeddings* exclusivamente a partir do recorte da região associada ao rosto contido nas imagens. Em seguida, é realizado o processo de clusterização, em que cada representação vetorial é associada a uma identidade específica. A Figura (4) apresenta os resultados simulados. As imagens autênticas (marcadas em vermelho) são agrupadas junto com as falsificações (marcadas em azul). Isto configura o desafio da distinção entre imagens autênticas e falsificações, uma vez ambas que estão contidas no mesmo agrupamento.

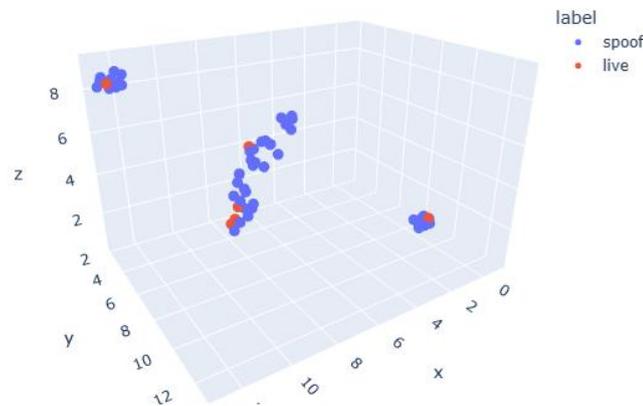


Figura 4. Resultado de clusterização conforme técnicas tradicionais de reconhecimento facial

Ao considerar a Figura (5), em que é apresentado um recorte facial empregado pelo algoritmo de reconhecimento na imagem associada a uma falsificação, observa-se a semelhança da imagem resultante a uma tentativa legítima. Sendo assim, uma representação vetorial gerada a partir de uma face legítima possivelmente não será distinta de uma falsificação. Este resultado reforça o desafio à distinção entre imagens autênticas e falsificadas, considerando que ambos os tipos são capazes de gerar representações vetoriais semelhantes e, conseqüentemente, são agrupados juntos no processo de clusterização.

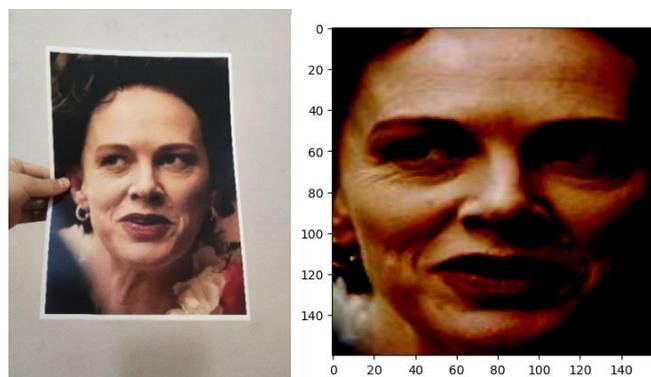


Figura 5. A imagem à esquerda representa uma tentativa de falsificação, enquanto a imagem à direita apresenta o rosto considerado

## 4.2 Método Proposto

Diferente dos métodos tradicionais de reconhecimento facial, o método proposto neste trabalho aplica o mecanismo de representação vetorial em toda a imagem, não se restringindo apenas à região do rosto. Este procedimento considera não apenas características faciais explicitamente identificáveis, mas também aspectos ambientais presentes na imagem, tais como iluminação, sombras, fundo e outros elementos contextuais que podem ser indicativos de tentativas de falsificação. Ao levar em conta esses elementos adicionais, a expectativa é que seja possível distinguir efetivamente as imagens que são tentativas de falsificação das imagens autênticas.

A Figura (6) apresenta o desempenho do método proposto neste trabalho. A técnica incorpora o uso do algoritmo t-SNE para a realização da clusterização em cima das representações vetoriais obtidas. No gráfico é possível observar como as imagens autênticas são eficientemente agrupadas, além disso, é possível realizar uma distinção das falsificações. Esse resultado demonstra o potencial de nosso método na melhoria do reconhecimento facial e na prevenção de falsificações, superando as limitações da abordagem tradicional.

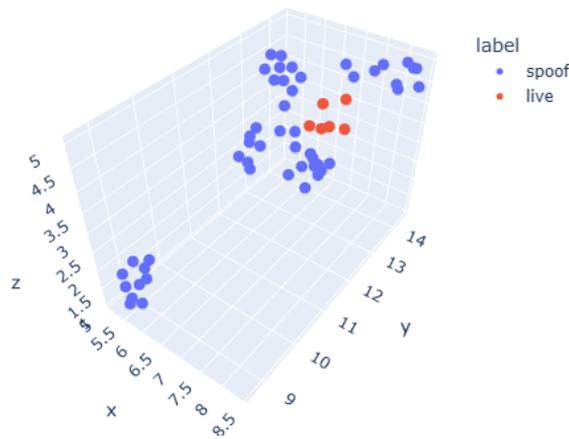


Figura 6. Clusterização conforme o método proposto

## 5. CONCLUSÃO

Este trabalho propôs um método inovador para prevenir falsificações em sistemas de reconhecimento facial. A abordagem emprega uma estratégia de clusterização em duas etapas, que aproveita técnicas avançadas de extração de *embeddings* e de redução de dimensionalidade, visando identificar e diferenciar entre imagens autênticas e falsificadas. Os resultados, apresentados na seção anterior, destacam as vantagens dessa abordagem, revelando uma capacidade promissora para melhorar a autenticação facial e prevenir falsificações.

Um aspecto essencial do nosso método é a utilização do algoritmo t-SNE para a etapa de clusterização. Este algoritmo provou ser eficiente na criação de grupos distintos de imagens autênticas e falsificadas, permitindo a diferenciação e a identificação precisa de falsificações. No entanto, embora o t-SNE tenha demonstrado um desempenho notável em nossos experimentos, é importante mencionar que outras técnicas de clusterização e redução de dimensionalidade poderiam ser empregadas conforme características específicas do problema e o contexto do conjunto de dados. A técnica de clusterização espectral, do inglês *Spectral Clustering* (Von Luxburg, et al. 2007), ou o algoritmo DBSCAN (Gholizadeh et al., 2021), poderiam oferecer outras perspectivas sobre a estrutura dos dados, possivelmente revelando padrões adicionais que podem ser úteis para a detecção de falsificações.

No que se refere à geração de *embeddings*, o nosso trabalho focou na utilização do modelo CLIP da OpenAI, contudo, outros modelos de geração de representações vetoriais, como *DeepFace* ou *FaceNet*, poderiam ser usados. Cada uma dessas ferramentas tem suas próprias características e pode ser adequada para diferentes contextos e tipos de dados.

Deste modo, o presente trabalho serve como uma base sólida para futuras pesquisas no campo de prevenção de falsificações em reconhecimento facial. Esperamos que o método proposto possa inspirar novas ideias e abordagens para combater eficientemente a falsificação e melhorar a autenticação facial. É importante mencionar que embora este trabalho se concentre no reconhecimento facial, o método proposto pode ser facilmente adaptado e aplicado a outros contextos em que a autenticação e a prevenção de falsificações são críticas. Acreditamos que a flexibilidade e a eficácia de nosso método têm o potencial de contribuir significativamente para diversos campos, incluindo a segurança digital e o controle de acesso.

## AGRADECIMENTOS

Agradeço o apoio técnico e computacional do LATITUDE, da UnB, que conta com apoio do CNPq - (Outorgas 312180/2019-5 PQ-2 e 465741/2014-2 INCT em Cibersegurança), da AGU (Outorga AGU 697.935/2019), ao Mestrado Profissional em Engenharia Elétrica, na área de concentração: Segurança Cibernética – 1ª Turma para Profissionais do Setor de Inteligência (Outorga ABIN 01/2019) ao DPI da UnB (Outorga 7129 FUB/EMENDA/DPI/COPEI/AMORIS) e do Projeto SISTER City (Outorga 625/2022) e a FAP/DF.

## REFERÊNCIAS

- Abdulrahman, M., Eleyan, A. (2015). Facial expression recognition using support vector machines. In: 2015 23rd signal processing and communications applications conference (SIU). IEEE, 2015. p. 276-279.
- Anthony, P., Ay, B., Aydin, G. (2021). A review of face anti-spoofing methods for face recognition systems. In: 2021 International Conference on INnovations in Intelligent SysTems and Applications (INISTA). IEEE, 2021. p. 1-9.
- Aria, Ms et al. (2020). Secure online payment with facial recognition using mtcnn. International Journal of Applied Engineering Research, v. 15, n. 3, p. 249-252, 2020.
- Dawson, M., Omar, M., Abramson, J. (2025). Understanding the methods behind cyber terrorism. In: Encyclopedia of Information Science and Technology, Third Edition. IGI Global, 2015. p. 1539-1549.
- Erdogmus, N., Marcel, S. (2013). Spoofing in 2d face recognition with 3d masks and anti-spoofing with kinect. In: 2013 IEEE sixth international conference on biometrics: theory, applications and systems (BTAS). IEEE, 2013. p. 1-6.
- Gholizadeh, N., Saadatfar, H., Hanafi, N. (2021). K-DBSCAN: An improved DBSCAN algorithm for big data. The Journal of supercomputing, v. 77, p. 6214-6235, 2021.
- Hoàng, G., Ngô, H. Face embedding. Disponível em: <https://hackmd.io/@gianghoangcotai/Sk05UiSFI>. Acesso em: 24 jul. 2023.
- Ikromovich, H. O. (2023). Mamatkulovich, Babakulov Bekzod. FACIAL RECOGNITION USING TRANSFER LEARNING IN THE DEEP CNN. Open Access Repository, v. 4, n. 3, p. 502-507, 2023.
- Kazemi, V., Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. p. 1867-1874.
- Kose, N., Dugelay, J.-L. (2013). On the vulnerability of face recognition systems to spoofing mask attacks. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, 2013. p. 2357-2361.
- Li, X., Yang, Z., Wu, H. (2020). Face detection based on receptive field enhanced multi-task cascaded convolutional neural networks. IEEE Access, v. 8, p. 174922-174930, 2020.
- Noble, W. S. (2006). What is a support vector machine?. Nature biotechnology, v. 24, n. 12, p. 1565-1567, 2006.
- Ochi, L. S., Dias, C. R., Soares, S. S. F. (2004). Clusterização em mineração de dados. Instituto de Computação-Universidade Federal Fluminense-Niterói, v. 1, p. 46, 2004.
- O'shea, K., Nash, R. (2007). An introduction to convolutional neural networks. arXiv preprint arXiv:1511.08458, 2015.
- Pan, G. et al. (2007). Eyeblink-based anti-spoofing in face recognition from a generic webcam. In: 2007 IEEE 11th international conference on computer vision. IEEE, 2007. p. 1-8.
- Parkin, A., Grinchuk, O. (2019). Recognizing multi-modal face spoofing with face recognition networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2019. p. 0-0.
- Pham, D. T., Dimov, S. S., Nguyen, C. D. (2005). Selection of K in K-means clustering. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, v. 219, n. 1, p. 103-119, 2005.

- Radford, A. et al. (2021). Learning transferable visual models from natural language supervision. In: International conference on machine learning. PMLR, 2021. p. 8748-8763.
- Sahoo, S. et al. (2023). The Universal NFT Vector Database: A Scaleable Vector Database for NFT Similarity Matching. arXiv preprint arXiv:2303.12998, 2023.
- Schroff, F., Kalenichenko, D., Philbin, J. (2015). Facenet: A unified embedding for face recognition and clustering. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2015. p. 815-823.
- Suwarno, S., Kevin, K. (2020). Analysis of face recognition algorithm: Dlib and opencv. Journal of Informatics and Telecommunication Engineering, v. 4, n. 1, p. 173-184, 2020.
- Taigman, Y. et al. (2014). Deepface: Closing the gap to human-level performance in face verification. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. p. 1701-1708.
- Van Der Maaten, L., and Hinton, G. (2008). "Visualizing data using t-SNE." Journal of machine learning research 9.11 (2008).
- Von Luxburg, U. (2007). A tutorial on spectral clustering. Statistics and computing, v. 17, p. 395-416, 2007.
- Xanthopoulos, P. et al. (2013). Linear discriminant analysis. Robust data mining, p. 27-33, 2013.
- Zhang, Y. et al. (2020). Celeba-spoof: Large-scale face anti-spoofing dataset with rich annotations. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII 16. Springer International Publishing, 2020. p. 70-85.