

Reconhecimento Facial com Anti-Spoofing

Desenvolvimento de Pipeline para Extração de Embeddings e
Detecção de Facial Landmarks

PEDRO RABELO MENDONÇA



UFG

UNIVERSIDADE
FEDERAL DE GOIÁS

UNIVERSIDADE FEDERAL DE GOIÁS (UFG)
INSTITUTO DE INFORMÁTICA (INF)

PEDRO RABELO MENDONÇA

Reconhecimento Facial com Anti-Spoofing

Desenvolvimento de Pipeline para Extração de Embeddings e Detecção de Facial
Landmarks

Goiânia
2025



UNIVERSIDADE FEDERAL DE GOIÁS
INSTITUTO DE INFORMÁTICA

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio do Repositório Institucional (RI/UFG), regulamentado pela Resolução CEPEC no 1240/2014, sem ressarcimento dos direitos autorais, de acordo com a Lei no 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo dos Trabalhos de Conclusão dos Cursos de Graduação disponibilizado no RI/UFG é de responsabilidade exclusiva dos autores. Ao encaminhar(em) o produto final, o(s) autor(a)(es)(as) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do Trabalho de Conclusão de Curso de Graduação (TCCG)

Nome(s) completo(s) do(a)(s) autor(a)(es)(as): PEDRO RABELO MENDONÇA

Título do trabalho: Reconhecimento Facial com Anti-Spoofing

Desenvolvimento de Pipeline para Extração de Embeddings e Detecção de Facial Landmarks

2. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador) Concorda com a liberação total do documento [X] SIM [] NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante: a) consulta ao(à)(s) autor(a)(es)(as) e ao(à) orientador(a); b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo do TCCG. O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro.

Obs.: Este termo deve ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Pedro Rabelo Mendonca, Discente**, em 04/02/2026, às 16:16, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Fernando Marques Federson, Professor do Magistério Superior**, em 13/03/2026, às 11:42, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5957103** e o código CRC **0F6093A4**.

Referência: Processo nº 23070.005552/2026-23

SEI nº 5957103

PEDRO RABELO MENDONÇA

Reconhecimento Facial com Anti-Spoofing

Desenvolvimento de Pipeline para Extração de Embeddings e Detecção de Facial Landmarks

Relatório final de Trabalho de Conclusão de Curso, apresentado à Universidade Federal de Goiás, como parte das exigências para a obtenção do título de Bacharel em Inteligência Artificial.

Orientador: Prof. Dr. Fernando Marques Federson

Goiânia

2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

MENDONÇA, PEDRO RABELO
Reconhecimento Facial com Anti-Spoofing [manuscrito]:
Desenvolvimento de Pipeline para Extração de Embeddings e Detecção de Facial
Landmarks / PEDRO RABELO MENDONÇA. - 2025.
56 f.: 2025

Orientador: Prof. Dr. Fernando Marques Federson
Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de
Goiás, Instituto de Informática (INF), Inteligência Artificial, Goiânia, 2025.

1. Inteligência Artificial. 2. Visão Computacional. 3. Reconhecimento
Facial.

I. Federson, Fernando Marques , orient. II. Título.

CDU 004

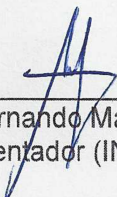
PEDRO RABELO MENDONÇA

Reconhecimento Facial com Anti-Spoofing


Desenvolvimento de Pipeline para Extração de Embeddings e Detecção de Facial Landmarks

Relatório final de Trabalho de Conclusão de Curso, apresentado à Universidade Federal de Goiás, como parte das exigências para a obtenção do título de Bacharel em Inteligência Artificial.


Data da Aprovação: 09 de dezembro de 2025.



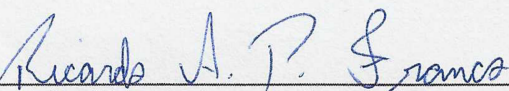
Prof. Dr. Fernando Marques Federson
Orientador (INF-UFG)



Prof. Dr. Aldo André Díaz Salazar
Coordenador de TCC do BIA (INF-UFG)



Prof. Dr. Anderson da Silva Soares
Coordenador do BIA (INF-UFG)



Prof. Dr. Ricardo Augusto Pereira Franco
(INF-UFG)

PEDRO RABELO MENDONÇA

Reconhecimento Facial com Anti-Spoofing

Desenvolvimento de Pipeline para Extração de Embeddings e Detecção de Facial Landmarks

RESUMO

Este Relatório de Conclusão de Curso tem como objetivo reunir os resultados da minha jornada para me tornar um especialista em **Reconhecimento Facial**. Uma ilustração e sua narrativa descrevem os períodos de trabalho. Os Apêndices contêm os Termos de Aceite de Entrega e os resultados obtidos durante cada período de trabalho.

Palavras-chave: Inteligência artificial; Visão computacional; Reconhecimento facial.

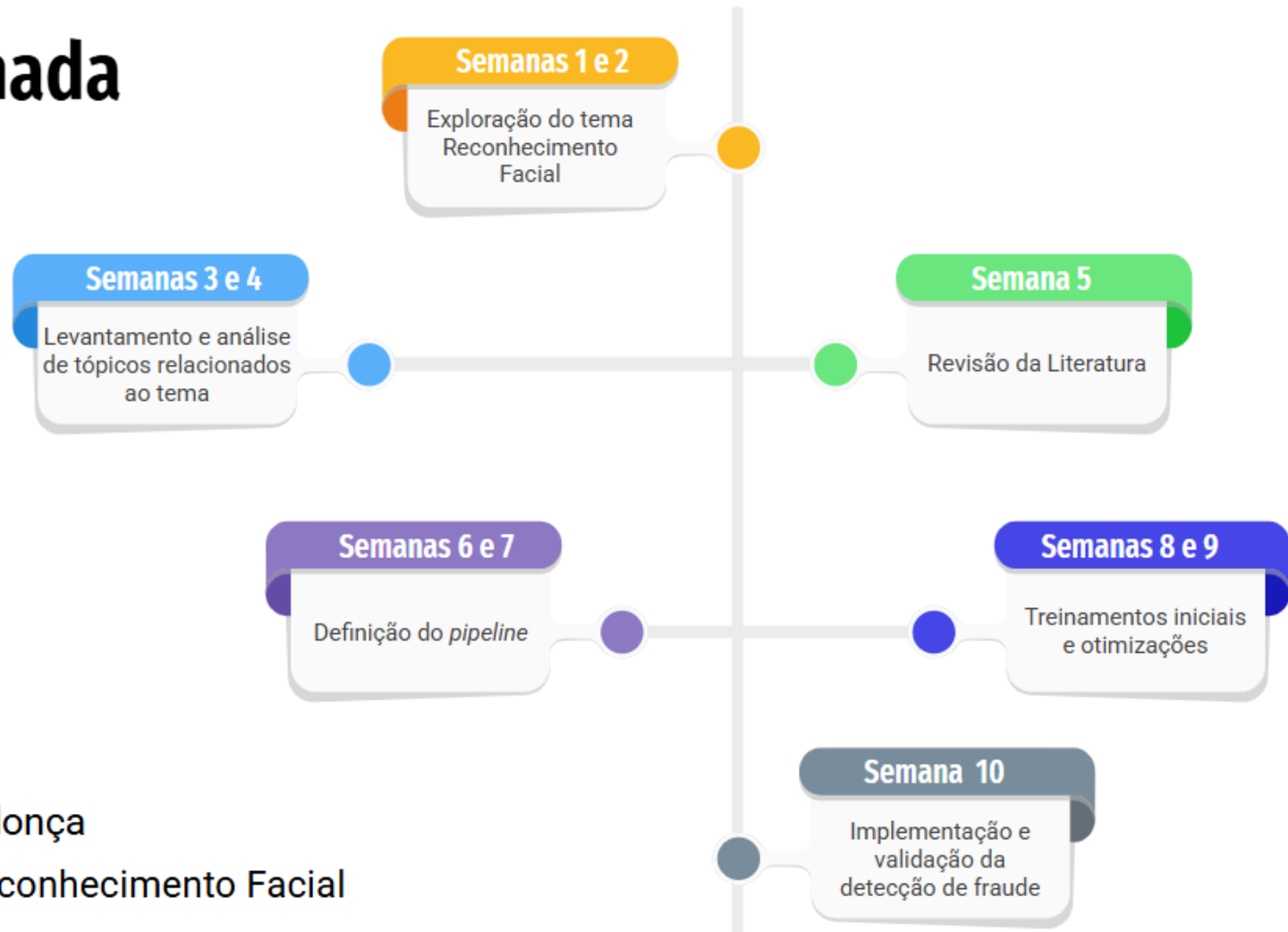
ABSTRACT

This Course Completion Report aims to bring together the results of my journey to become an expert in **Facial Recognition**. An illustration and its narrative describe the work periods. The Appendices contain the Delivery Acceptance Terms and the results obtained during each work period.

Keywords: Artificial intelligence; Computer vision; Facial recognition.

Goiânia
2025

Minha Jornada



Pedro Rabelo Mendonça
Especialista em: Reconhecimento Facial

MINHA JORNADA

Nome: Pedro Rabelo Mendonça

Especialidade: Reconhecimento Facial

Objetivo deste documento

Durante o processo da disciplina Residência em IA¹, foram gerados diversos resultados na construção da minha especialização. A cada semana, um conjunto de resultados foi formalizado por um Termo de Aceite de Entrega e avaliado por uma banca, considerando o planejado e o realizado para o período. Este documento tem como objetivo descrever esses resultados obtidos, fazendo referência aos Termos de Aceite de Entrega e seus documentos associados.

Minha Jornada

O início da minha jornada se deu com as **Semanas 1 e 2**, nas quais, embora já houvesse definido o interesse por Visão Computacional, mais especificamente Reconhecimento Facial, percebi que abordar diretamente um tópico específico não contribuiria para uma formação completa e especializada da forma desejada. Assim, iniciei estudando os níveis de abstração do tema e conduzi uma série de pesquisas bibliográficas de artigos, cujos resultados podem ser observados nos **Apêndices 1**. Inicialmente, realizei uma pesquisa exploratória utilizando como filtro apenas o período de publicação; posteriormente, adotei uma abordagem mais criteriosa por meio da plataforma Parsifal para complementar minha base de artigos. Esta etapa foi crucial não apenas para fornecer uma visão ampla do campo e possibilitar uma escolha mais apropriada do escopo do trabalho, como também para estabelecer uma base bibliográfica sólida que fundamentaria as etapas subsequentes.

¹ Dez Semanas, entre setembro de 2025 e dezembro de 2025.

Nas **Semanas 3 e 4**, o foco principal consistiu em realizar um levantamento bibliográfico de forma detalhada e aprofundada para analisar os tópicos englobados pelo Reconhecimento Facial. Todos os artigos levantados foram devidamente organizados e categorizados em pastas de acordo com seus respectivos assuntos principais, conforme documentado no **Apêndice 2**. Durante este processo, uma publicação que contribuiu significativamente para a compreensão da área foi o artigo de Phillips et al.², que apresenta uma revisão abrangente dos 50 anos de desenvolvimento do Reconhecimento Facial automatizado. Nesta publicação, obtive uma visão ampla do tema, abordando diversos tópicos do passado, presente e expectativas para o futuro, incluindo: história da área, principais avanços tecnológicos, desafios atuais, uso de dados sintéticos, arquiteturas de modelos, métodos de avaliação e benchmarks, direções futuras, além de questões sobre explicabilidade e interpretabilidade dos sistemas. O resumo detalhado deste trabalho encontra-se no documento "Resumo Artigos" presente no **Apêndice 2**. Além disso, realizei uma série de estudos envolvendo conceitos e teorias fundamentais utilizados não apenas em Reconhecimento Facial, mas também em Visão Computacional de forma mais ampla. Este processo foi conduzido por considerar de suma importância revisar e dominar tais fundamentos antes de avançar para a especialização na área. Estas revisões conceituais foram documentadas no arquivo "Conhecimentos Base", também disponível no **Apêndice 2**. Ao término desta etapa, pude obter maior clareza sobre os principais tópicos da área, possibilitando uma tomada de decisão mais fundamentada e consciente.

Dando prosseguimento à jornada, na **Semana 5** continuei a revisão para completar e aprofundar o estudo dos artigos previamente selecionados. Nesta etapa, foram elaborados novos resumos analíticos dos trabalhos escolhidos e foi realizada a atualização do documento "Conhecimentos Base" com novos conceitos e informações relevantes, disponíveis. Além disso, foi realizada uma nova rodada de seleção de artigos para complementar a base bibliográfica existente. Para esta tarefa, utilizei como ferramenta de auxílio o Claude Sonnet 4.5, empregando "prompts" detalhadamente construídos para identificar e recuperar artigos relevantes, atuais e alinhados ao estado da arte da área. Ao

² Phillips, P. J., Yates, A. N., Hu, Y., Hahn, C. A., Noyes, E., Jackson, K., ... & O'Toole, A. J. (2025). *50 Years of Automated Face Recognition*. Michigan State University.

final deste processo, foram coletados 14 artigos, selecionados após criteriosa análise que incluiu a leitura prévia e o descarte daqueles que não atendiam aos critérios de relevância estabelecidos. Todo este processo de seleção e triagem está documentado no **Apêndice 2** e no **Apêndice 3**.

Durante as **Semanas 6 e 7**, o foco direcionou-se ao desenvolvimento de um *pipeline* completo capaz de realizar o pré-processamento de imagens, treinamento do modelo e inferência. O pré-processamento consiste em: detecção facial, alinhamento facial, crop para isolar apenas a região da face e anotação dos 5 *facial landmarks* para utilização posterior. O treinamento, por sua vez, empregaria uma rede neural profunda com três cabeças: a primeira para extração de *embeddings*, a segunda para detecção dos 5 *facial landmarks* e a terceira para classificação binária de *anti-spoofing*, determinando se uma imagem corresponde a uma pessoa real ou a uma tentativa de fraude, prevenindo assim ataques de apresentação. Inicialmente, foram realizados testes com o MTCNN³ (*Multi-task Cascaded Convolutional Networks*) para compreender melhor seu funcionamento e as saídas geradas. Nestes experimentos, a face e os 5 *facial landmarks* (centro dos dois olhos, ponta do nariz e cantos da boca) foram detectados em imagens de teste para melhor visualização, cujos resultados podem ser observados no **Apêndice 4**. Outra referência fundamental nesta etapa foi o trabalho de Melgor⁴ sobre pré-processamento facial. Posteriormente, foram definidos os componentes principais do modelo: a arquitetura MobileNetV3 Large como *backbone*, escolhida em virtude de seu excelente equilíbrio entre desempenho e eficiência computacional, e a função de perda CosFace, uma das mais utilizadas na literatura com resultados expressivos, conforme observado nos artigos revisados. Com estas definições, deu-se início à implementação do pipeline descrito.

Com foco maior em testes e otimizações, durante as **Semanas 8 e 9**, foi finalizada a primeira versão do *pipeline*, conforme descrito, porém ainda sem a implementação do módulo de *anti-spoofing*. Para todos os experimentos subsequentes, foram utilizados os

³ Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks. *IEEE Signal Processing Letters*, 23(10), 1499-1503.

⁴ Melgor. (2017). *Demystifying Face Recognition III: Face Preprocessing*. Disponível em: <https://melgor.github.io/blcv.github.io/static/2017/12/28/demystifying-face-recognition-iii-face-preprocessing/index.html>

seguintes *datasets*: VGGFace2 112x112 para treinamento e validação, e LFW (*Labeled Faces in the Wild*) para teste. Uma observação importante refere-se ao fato de que, devido a limitações de capacidade computacional, foi utilizada apenas uma subamostragem do VGGFace2, correspondente a 90 identidades (de um total de 8.631) e 31.279 imagens (de um total de 3.310.000). No primeiro teste, os resultados foram considerados satisfatórios em função da quantidade de dados utilizados, alcançando acurácia de 92,17% no conjunto de validação e 65,95% no LFW. O processo completo e as demais métricas deste primeiro experimento podem ser consultados no documento "Teste Pipeline", presente no **Apêndice 5**. No segundo treinamento, pequenas melhorias foram implementadas na etapa de pré-processamento, resultando em otimização dos resultados anteriores: a acurácia no LFW aumentou para 69,10%. É importante ressaltar que, com exceção das alterações no pré-processamento, todas as demais condições e hiperparâmetros de treinamento permaneceram idênticos entre os dois experimentos. Para informações detalhadas, consulte o documento "Teste Pipeline 2" no **Apêndice 5**.

Na **Semana 10**, o foco concentrou-se na implementação e validação do módulo de *anti-spoofing* para detecção de fraudes e ataques de apresentação. Para o treinamento desta terceira cabeça da rede, foi utilizado o *dataset* CASIA-FASD (*Chinese Academy of Sciences' Institute of Automation Face Anti-Spoofing Database*), contendo 115.494 amostras. A adoção de um novo *dataset* justifica-se pela necessidade de anotações específicas para detecção de ataques de apresentação, diferentes das disponíveis nos conjuntos de dados anteriormente utilizados, uma vez que o objetivo desta camada difere substancialmente das demais. Para validar o *pipeline* completo, foram estabelecidos três cenários críticos para verificar a capacidade do modelo em: (1) reconhecer a mesma pessoa a partir de duas imagens diferentes; (2) diferenciar corretamente duas pessoas; e (3) detectar tentativas de fraude, nas quais uma pessoa tenta se passar por outra. O modelo demonstrou desempenho satisfatório nos três cenários, identificando corretamente cada situação proposta. Todo o processo de treinamento, bem como os experimentos de validação e seus resultados, estão documentados no arquivo "Anti-Spoofing", disponível no **Apêndice 6**.

Expresso meus agradecimentos ao programa de Residência em IA pela oportunidade de desenvolvimento, ao orientador responsável pelo acompanhamento técnico e aos demais participantes pelas contribuições durante o processo. A experiência adquirida nestas **Semanas** foi determinante para o aprofundamento nos fundamentos teóricos e práticos de Reconhecimento Facial e Visão Computacional.

APÊNDICE 1

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 3 de set. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Foram realizadas as seguintes atividades:

- Primeiramente foi feita uma pesquisa literária abrangente para compreender melhor o cenário atual do Reconhecimento Facial, com isso, foi possível destacar os seguintes principais pontos:
 - CNNs x ViTs
 - Anti-spoofing e Liveness detection
 - Dificuldades presentes nas aplicações práticas do mundo real (objetos de oclusão, iluminação, qualidade da imagem e variações de ângulo e posição)
 - Técnicas multimodais de verificação biométrica
 - Viés demográfico
 - Evolução das diferentes técnicas utilizadas para reconhecimento facial
 - [Pesquisa Literária FR](#)
- Definição do tema central a ser abordado na Residência a partir da pesquisa anterior:
 - Transição de Arquiteturas Convolucionais para Vision Transformers em Reconhecimento Facial
- Dado a definição do escopo, fiz uma segunda pesquisa para filtrar melhor os artigos que se encaixam dentro da abordagem adotada
 - Artigo que compara as duas arquiteturas aplicadas em reconhecimento facial
 - Artigos que abordam especificamente os avanços gerados pelo uso dos ViTs
 - Artigos estado da arte em reconhecimento facial com CNNs
 - Benchmark MFR-Ongoing Challenge
 - [Artigos de apoio para estudo da transição de CNNxViT](#)

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Para a próxima Semana, pretendo realizar as seguintes tarefas:

- Aprofundamento nos artigos finais selecionados (artigos presentes no arquivo: Artigos de apoio para estudo da transição de CNNxViT)
- Estudo das arquiteturas e códigos (se disponíveis) de cada artigo para futura aplicação e realização de testes

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: Go! ▾

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 11 de set. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Durante essa Semana, foram desenvolvidas as seguintes atividades

- Partindo do tópico específico selecionado na Semana anterior, fiz um estudo para subir os níveis de abstração e conseguir escolher um tema de forma mais apropriada para a Residência
 - Foram utilizadas ferramentas de LLM como Claude Pro e Gemini 2.5 Pro para verificação dos níveis de abstração;
 - Escolha do tema: Estudo do Reconhecimento Facial no campo da Visão Computacional.
 - **Níveis de Abstração**
- Dado a escolha do tema, foi feita uma pesquisa abrangente para artigos e materiais sobre com o objetivo de juntar uma base literária sólida. Para isso, foi utilizado o Parsifal, permitindo definir os principais termos da pesquisa, como:
 - Objetivo
 - Palavras chaves
 - Palavras de busca
 - Fontes
 - **Pesquisa**
- Principais assuntos da pesquisa:
 - Abordagens Metodológicas
 - Arquiteturas das Redes Neurais
 - Tipos de Tarefa (identificação facial, verificação facial, ...)
 - Desafios Técnicos
 - Avaliações (benchmarks) e Datasets
- Papers coletados ao todo
 - 93

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

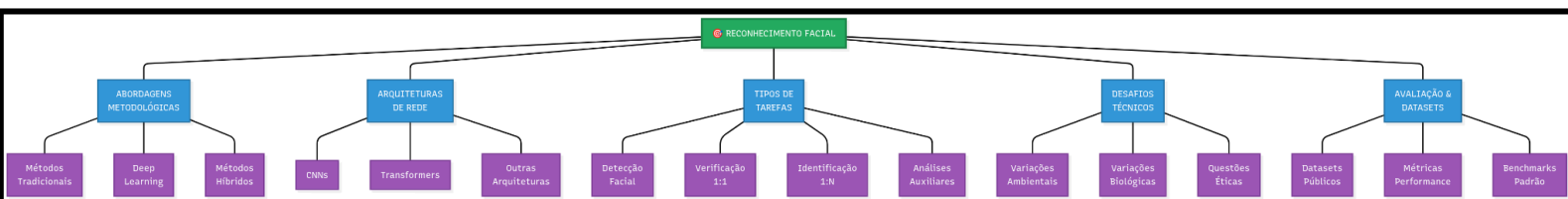
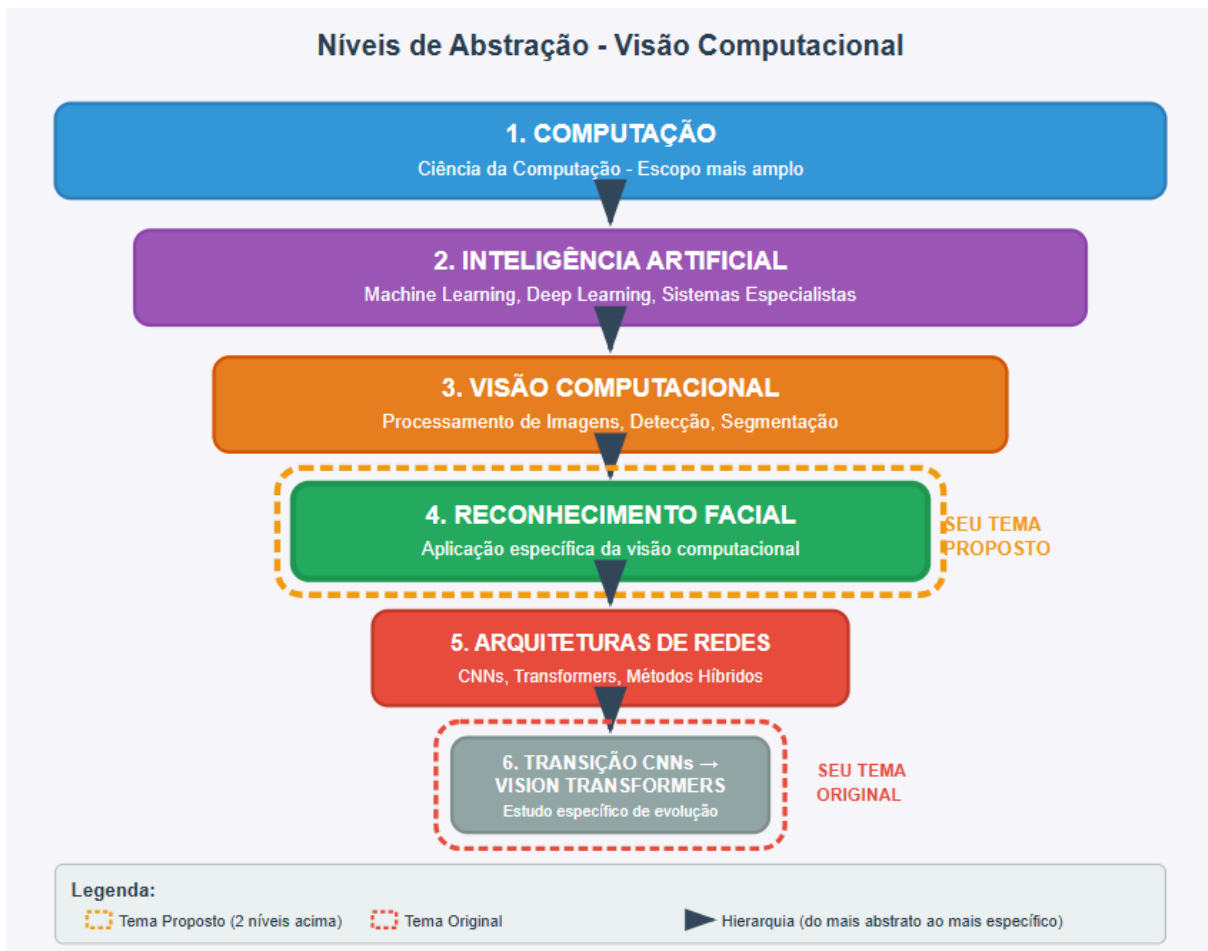
Assim, para a próxima Semana, planejo filtrar os artigos coletados, ainda utilizando a plataforma Parsifal como suporte, e organizar eles dentro dos 5 assuntos definidos.

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: Go! ▾

Níveis de Abstração



Resumo

Este documento descreve o processo metodológico de definição do escopo da pesquisa por meio da análise de níveis de abstração, adotando uma abordagem *bottom-up* para determinar o tema central da Residência. Partindo inicialmente do tópico específico sobre a transição entre Arquiteturas Convolucionais e *Vision Transformers*, e com o suporte de ferramentas de Modelos de Linguagem (LLMs) para validação conceitual e estruturação visual, elevou-se o nível de organização para definir o tema "Estudo do Reconhecimento Facial no campo da Visão Computacional". A seleção deste nível de abstração justifica-se pela necessidade de uma perspectiva mais abrangente, capaz de integrar diversas arquiteturas de redes neurais, abordagens metodológicas, desafios técnicos e bases de dados, resultando na elaboração de diagramas estruturais que orientarão o prosseguimento da investigação científica.

Pesquisa e Pesquisa Literária FR

Resumo

A presente fundamentação metodológica e teórica delinea o escopo da investigação científica proposta, integrando uma Revisão Sistemática da Literatura (RSL) com uma pesquisa bibliográfica exploratória sobre o estado da arte em Reconhecimento Facial. Para a execução da RSL, utilizou-se a plataforma Parsifal como ferramenta de gestão, visando analisar o período compreendido entre 2020 e 2025. O protocolo de pesquisa foi estruturado a partir de questões norteadoras focadas em identificar abordagens metodológicas, comparar arquiteturas de redes neurais (CNNs e *Vision Transformers*), mapear desafios técnicos e examinar métricas de avaliação. A estratégia de busca empregou *strings* booleanas contendo termos essenciais como "*face recognition*" e "*deep learning*", aplicadas em bases de dados de relevância acadêmica como ACM Digital Library, IEEE Xplore, ScienceDirect e arXiv. Após a aplicação rigorosa de critérios de inclusão e exclusão, selecionou-se um *corpus* final de 93 artigos para compor o estudo aprofundado da Residência.

Paralelamente, a pesquisa bibliográfica preliminar, essencial para o refinamento do escopo do estudo, traçou a evolução histórica da área, desde métodos geométricos iniciais até o atual paradigma de *Deep Learning*. A análise da literatura indicou que, embora a tecnologia de reconhecimento facial esteja consolidada com acurácias frequentemente superiores a 99%, persistem desafios significativos em cenários não controlados, envolvendo variações de iluminação e oclusão. Destaca-se, na literatura recente, uma transição tecnológica emergente: estudos comparativos demonstram que arquiteturas baseadas em *Vision Transformers* têm superado as tradicionais Redes Neurais Convolucionais (CNNs) em quesitos críticos como acurácia, robustez a oclusões e tempo de inferência.

Por fim, a síntese dos trabalhos analisados revelou lacunas importantes que justificam a continuidade da investigação. Embora as CNNs ainda dominem o cenário, foram identificadas limitações quanto à representatividade dos *datasets* — que frequentemente introduzem vieses demográficos — e à inadequação de certas métricas de avaliação frente às complexidades do mundo real. Ademais, questões éticas relacionadas à privacidade e discriminação algorítmica emergem como pontos críticos. Diante desse contexto, o presente trabalho se propõe a investigar essas lacunas, focando no desenvolvimento de sistemas mais justos, na análise de abordagens multimodais e na avaliação de novas arquiteturas que possam mitigar os problemas de robustez e viés identificados no estado da arte atual

APÊNDICE 2

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 18 de set. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Dando prosseguimento ao Tema: Estudo do Reconhecimento Facial (Visão Computacional). Nesta Semana, foram realizadas as seguintes atividades:

- Foi dada continuidade à pesquisa de base literária que estava sendo feita no Parsifal:
 - Leitura para conseguir ter uma visão geral dos 93 artigos coletados para aprovar/rejeitar eles baseando-se no critério de conformidade do tema geral e data
 - Dado o passo acima, sobram 12 artigos
 - No qual apenas 5 estavam disponíveis
- Complementação dos artigos coletados pelos artigos já selecionados previamente na primeira pesquisa do primeiro Stage.
 - Pesquisa Literária FR
 - Artigos de apoio para estudo da transição de CNNxViT
- Organização dos artigos resultantes em pastas de acordo com cada um dos subtemas já definidos (com a adição de: História)
 - História
 - Abordagens Metodológicas
 - Arquiteturas
 - Tipos de Tarefa
 - Desafios Técnicos
 - Benchmarks e Datasets

Link com a documentação do que foi realizado acima: Pesquisa pt.2

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Para a próxima Semana, planeje o seguinte:

- Iniciar as leituras e estudos dos artigos, primeiramente da pasta: História
- Gerar um material em cima do que foi lido

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: Go! ▾

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 25 de set. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA


Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Tema Geral da minha Residência: Reconhecimento Facial

Durante esta Semana realizei as seguintes tarefas:

- presentes. Organizei por títulos e fiz em um documento no qual pretendo ir adicionando os demais resumos de artigos que for lendo daqui para frente.
 - [Resumo Artigos](#)
- Com isso, pude ter uma clareza maior sobre qual seria de fato o tema que eu iria atacar na minha Residência: **Treinamento de Redes Neurais para Reconhecimento Facial**.
- As motivações para este tema são:
 - Durante a leitura do artigo, ele deixa bem claro que não existe o melhor modelo para todas as tarefas de reconhecimento facial;
 - Outro ponto foi justamente a estatística apresentada no artigo de que o maior número de papers nesta área são sobre Funções de Loss;
 - Dado essas informações, me senti atraído a ideia de me especializar em uma área em que há bastante conhecimento para ser explorado e que cada situação precisa ser analisada (por um especialista rs) para determinar o conjunto correto a ser aplicado.
- Primeiramente realizei a leitura do artigo “50 Years of Automated Face Recognition” de Michigan State University, publicado em 30 de Maio de 2025 e possuindo 20 páginas.
- Sua leitura foi realizada por abordar o desenvolvimento da área de Reconhecimento Facial durante os últimos 50 anos e abordando suas principais áreas.
- Após sua leitura, foi feito um resumo sintetizando seus conteúdos nos principais tópicos
- Meus objetivos com esse tema são me especializar nos seguintes ramos presentes no processo de treinamento de redes neurais para Reconhecimento Facial:
 - Técnicas de pré-processamento de imagens
 - Arquiteturas
 - Backbones
 - Funções de Loss
- Posteriormente, decidi fazer uma segunda leitura no artigo para ir construindo um documento como uma espécie de “base de conhecimento” com todos os termos gerais e até mesmo curiosidades que acredito que vão ser importantes para minha formação.
- Busquei explorar termos e referências externas presentes no artigo para a partir deles conseguir

uma visão geral daquele assunto.

- Esse documento vai estar em construção até a finalização do meu processo na Residência.
 -  Conhecimentos Base

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Para a próxima Semana, planejo:

- Dar continuidade a construção do meu documento “Conhecimento Base”
- Realizar a leitura do artigo “Facial Recognition Algorithms: A Systematic Literature Review”
- Deixar de lado artigos que fogem do meu escopo
- Utilizar essas duas fontes e outros meios de pesquisa para selecionar artigos mais direcionados ao meu tema selecionado
- Iniciar a leitura destes últimos artigos selecionados e os anteriormente selecionados

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

Essa Semana fiquei doente e por causa disso tive pequenos atrasos no meu desenvolvimento

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: 

Resumo Artigos

Resumo

O documento analisa a trajetória histórica e tecnológica do reconhecimento facial ao longo dos últimos 50 anos, delineando a evolução desde os primeiros sistemas baseados em métodos geométricos e identificação manual de marcos (1966-1973) até o atual paradigma de *Deep Learning*. A literatura destaca que, embora as Redes Neurais Convolucionais (CNNs) tenham revolucionado a área ao permitir o aprendizado de representações discriminativas em grandes *datasets*, observa-se atualmente uma transição tecnológica emergente para arquiteturas de *Vision Transformers* (ViT), que oferecem maior capacidade de capturar dependências globais nas imagens. Uma revisão sistemática abrangendo o período recente reforça a supremacia do *Deep Learning* na melhoria da precisão e eficiência, mas alerta para a estagnação das métricas de avaliação tradicionais (como acurácia e F1 score), que frequentemente falham em capturar a complexidade dos cenários do mundo real.

No âmbito das arquiteturas específicas de detecção e alinhamento, o texto detalha o funcionamento do MTCNN e do RetinaFace como soluções fundamentais. O MTCNN é descrito como um *framework* que integra detecção e alinhamento através de uma arquitetura em cascata de três estágios (P-Net, R-Net, O-Net), utilizando a estratégia de *online hard sample mining* para otimizar o treinamento e eliminar a seleção manual de amostras difíceis. Avançando na complexidade, o RetinaFace é apresentado como um detector *single-stage* que realiza localização facial densa, incorporando simultaneamente detecção, alinhamento, segmentação *pixel-wise* e regressão de correspondência 3D densa. Diferentemente de seus predecessores, o RetinaFace inova ao combinar aprendizado supervisionado com auto-supervisionado (*self-supervised*) para a malha 3D, alcançando resultados do estado da arte em *benchmarks* desafiadores como o WIDER FACE.

Por fim, o estudo levanta questões críticas sobre as limitações éticas e técnicas que persistem na área. Identificou-se que os *datasets* mais utilizados, como LFW e CASIA WebFace, sofrem de falta de diversidade demográfica e geográfica, introduzindo vieses algorítmicos significativos e prejudicando a generalização dos modelos. O documento enfatiza a necessidade urgente de desenvolver modelos explicáveis (*Explainable AI*) e sistemas multimodais para aumentar a transparência e a robustez contra ataques de *spoofing*. As direções futuras apontam para o uso de dados sintéticos para mitigar problemas de privacidade e escassez de dados, bem como a integração de modelos fundacionais (como CLIP e DINOv2) para aprimorar a representação de características em cenários com poucos dados.

Referências

- MICHIGAN STATE UNIVERSITY. **50 Years of Automated Face Recognition**. 30 maio 2025.
- EL FADEL, Nazar. Facial Recognition Algorithms: A Systematic Literature Review. **Journal of Imaging**, 13 fev. 2025.
- SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY; CHINESE ACADEMY OF SCIENCES; THE CHINESE UNIVERSITY OF HONG KONG. **Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks**. 2016.
- IMPERIAL COLLEGE LONDON; INSIGHTFACE. **RetinaFace - Single-stage Dense Face Localisation in the Wild**. 04 maio 2019.

Conhecimentos Base

Resumo

O presente documento estabelece uma fundamentação teórica abrangente sobre o estado da arte do reconhecimento facial, traçando sua evolução histórica desde os sistemas pioneiros de Takeo Kanade em 1973 até o domínio atual das técnicas de *Deep Learning*. O texto define a taxonomia operacional dos sistemas, dividindo-os em tarefas de verificação (1:1) e identificação (1:N), e destaca o dilema fundamental entre a alta performance discriminativa dos modelos modernos e a sua falta de interpretabilidade, resultando em "caixas-pretas" que dificultam a explicabilidade das decisões. Além disso, enfatiza-se a importância crítica do pré-processamento, citando o MTCNN e o *RetinaFace* como paradigmas essenciais para a detecção e alinhamento facial, sendo o último destacado pelo uso de redes piramidais de características e aprendizado multitarefa.

No âmbito das arquiteturas de redes neurais, o documento descreve a consolidação das Redes Neurais Convolucionais (CNNs), especificamente as variantes ResNet e IR-ResNet, que dominaram o cenário na última década. Contudo, aponta para uma mudança de paradigma emergente com a ascensão dos *Vision Transformers* (ViTs), que processam imagens como sequências de *patches* para capturar dependências globais. São detalhadas adaptações específicas para faces, como o TransFace e o LVFace (estado da arte em 2025), este último integrando a codificação de posição relativa baseada em *landmarks* para superar as limitações dos modelos anteriores. A evolução dos *datasets* também é discutida, demonstrando um crescimento exponencial desde o CASIA-WebFace até bases massivas como o WebFace42M, essenciais para o treinamento dessas arquiteturas complexas.

A eficácia dos sistemas de reconhecimento facial é atribuída, em grande parte, à evolução das funções de perda (*Loss Functions*). O texto detalha a transição de abordagens baseadas em *Contrastive Learning*, como a *Triplet Loss*, para as dominantes *Margin-based*

Softmax Losses. Métodos como ArcFace e CosFace são explicados pela sua capacidade de maximizar a separação inter-classe através de margens angulares e aditivas. Avanços recentes incluem margens adaptativas, como o AdaFace, que ajustam a margem dinamicamente com base na qualidade da imagem e na dificuldade da amostra, oferecendo maior robustez em *benchmarks* desafiadores como o IJB-C.

Por fim, o documento aborda as práticas de engenharia necessárias para a implantação em produção e as considerações éticas indispensáveis. São descritas técnicas de otimização como quantização e destilação de conhecimento (modelo professor-aluno) para viabilizar a execução em dispositivos com recursos limitados. Do ponto de vista ético, discute-se a necessidade de métricas de justiça (*Fairness*) para mitigar vieses demográficos e a implementação de proteções de privacidade. As tendências futuras apontadas incluem o uso de dados sintéticos para superar barreiras de privacidade, a adoção de modelos fundacionais de larga escala e o desenvolvimento de arquiteturas híbridas que unem a eficiência das CNNs com a capacidade global dos ViTs.

APÊNDICE 3

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 1 de out. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Durante esta Semana, foram realizadas as seguintes atividades:

- Tema: **Treinamento de Redes Neurais para Reconhecimento Facial.**
- Foi dado continuidade ao desenvolvimento dos meus resumos, desta vez, sobre o artigo “**Facial Recognition Algorithms: A Systematic Literature Review**”, do Journal of Imaging publicado em 13 de fevereiro de 2025 contendo 29 páginas.
 - [Resumo Artigos](#)
- Foram acrescentadas mais informações dentro do meu documento de conhecimento base (data augmentation, benchmarks, novas arquiteturas e conhecimentos mais gerais de visão).
 - [Conhecimentos Base](#)
- Agregação de artigos mais focados no meu tema, utilizando Claude Sonnet 4.5 como ferramenta de auxílio. Organização dos mesmos em pastas no drive.
 - [Complementação de Artigos](#)

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Para a próxima Semana, planejo:

- Ler e resumir os artigos de pré processamento selecionados
- Documentar principais estratégias e técnicas utilizadas
- Implementá-las

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: [Go! ▾](#)

Complementação Artigos

Resumo

Este documento descreve uma estratégia detalhada de levantamento bibliográfico, auxiliada por ferramentas de LLM, com o objetivo de identificar os artigos mais fundamentais e relevantes no campo do reconhecimento facial. A pesquisa foi estruturada em cinco categorias principais: pré-processamento de imagens, arquiteturas CNN, arquiteturas Vision Transformers (ViT), modelos híbridos e funções de perda, priorizando trabalhos com alto desempenho em benchmarks e elevada contagem de citações. Para o pré-processamento, foram selecionados o *RetinaFace* e o *MTCNN* como padrões de indústria para detecção e alinhamento, além de estratégias de *data augmentation* essenciais para o treinamento robusto.

No que tange às arquiteturas de redes neurais, o levantamento consolidou a importância das CNNs clássicas, destacando o *ResNet* e o *VGGFace* como bases históricas, e o *ArcFace* (utilizando IR-ResNet) como o padrão atual. Paralelamente, identificou-se a emergência dos *Vision Transformers*, citando o *ViT* original e adaptações específicas para faces como *TransFace* e *LVFace*, que oferecem maior robustez a oclusões e variações de distância. Observou-se que a literatura sobre arquiteturas híbridas ainda é difusa, recomendando-se focar na comparação entre os paradigmas puros de CNN e ViT.

Por fim, o documento traça a evolução crítica das funções de perda (*Loss Functions*), partindo de abordagens de aprendizado métrico como o *FaceNet (Triplet Loss)* e *Center Loss*. A análise progride para as perdas baseadas em margem angular, incluindo *SphereFace*, *CosFace* e *ArcFace*, que melhoraram significativamente a discriminação intra e inter-classe. O levantamento culmina com métodos de margem adaptativa, como o *AdaFace*, que ajustam a margem com base na qualidade da imagem, resultando em uma

coleção organizada de artigos que fundamentam o conhecimento técnico necessário para a pesquisa.

APÊNDICE 4

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 8 de out. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Durante essa Semana, foram realizadas as seguintes tarefas em relação ao tema **Reconhecimento Facial**:

- Leitura e realização do resumo do artigo “Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks” (MTCNN).
- Leitura e realização do resumo do artigo “RetinaFace - Single-stage Dense Face Localisation in the Wild”.
 - [Resumo Artigos](#)
- Após a leitura dos artigos, fui atrás de investigar alguns pontos importantes do pré-processamento aplicado atualmente para imagens de Reconhecimento Facial. Com isso, descobri os seguintes pontos:
 - Uso de 5 facial landmarks
 - 2D Alignment > 3D Alignment
 - Crop é muito importante
 - Referência: [Demystifying Face Recognition IV: Face-Alignment](#)
- Descoberta de método mais atual e com melhores resultados: “Sample and Computation Redistribution for Face Detection”.
- Assim, foi iniciada a implementação do MTCNN.
 - [GitHub](#)

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Para a próxima Semana, continuar o desenvolvimento, fazer uma análise dos resultados do MTCNN e me aprofundar ainda mais nele, procurando pontos de ajustes.

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: [Go!](#)

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 16 de out. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Durante esta Semana, foi dado prosseguimento ao desenvolvimento das atividades relacionadas ao tema **Reconhecimento Facial**, sendo realizadas as seguintes atividades:

- Implementação da arquitetura do MTCNN.
 - [GitHub](#)
- Notebook para demonstração de inferência.
 - [MTCNN.ipynb](#)
- Com isso, foram realizados alguns testes para garantir que estava funcionando corretamente (por funcionar corretamente = detectar as faces e identificar o 5 landmarks).
 - [Resultados](#)
- A partir dos resultados obtidos, decidi criar um pipeline completo que abordasse tanto as técnicas já presentes no MTCNN, como também outras que pudessem ser aplicadas no próprio treinamento da rede. Dessa forma, dei início a seguinte implementação que segue os passos:
 - MTCNN (detecção da face, crop e face alignment)
 - MobileNetV3 Large (backbone escolhido)
 - CosFace (loss escolhida)
 - Multi-task Learning com os 5 landmarks

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: [Go! ▾](#)

Resultados

Resumo

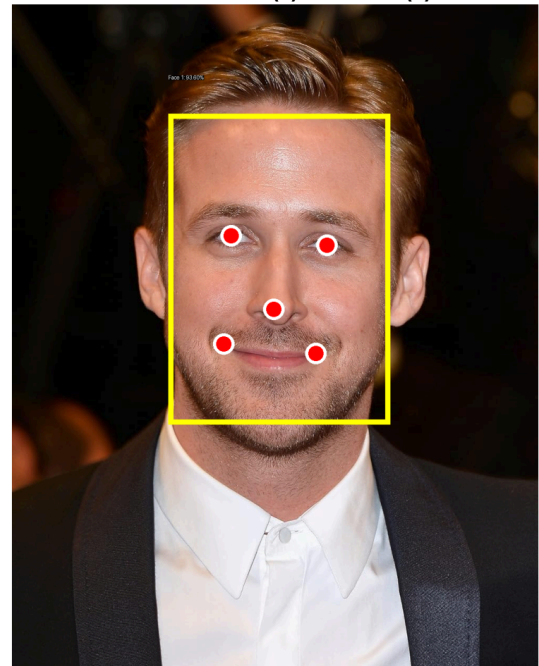
Neste documento encontram-se os primeiros resultados e experimentos feitos com o MTCNN para maior familiarização da ferramenta. São descritos testes feitos em algumas imagens e os valores que são retornados em cada uma, dentro delas: nível de confiança de detecção de cada face, coordenadas do bounding box e coordenadas dos pontos faciais (facial landmarks).

Exemplo

ANTES - Imagem Original



DEPOIS - 1 face(s) detectada(s)



DETAHES DAS DETECCÖES

Face 1:

Confiança: 0.9360 (93.60%)

Bounding Box: [297, 208, 421, 591]

Landmarks: ['nose', 'mouth right', 'right eye', 'left eye',
'mouth left']

APÊNDICE 5

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 23 de out. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Durante as Semanas anteriores, foram feitas revisões literárias sobre o tema escolhido que foi **Reconhecimento Facial**. Posteriormente, aprofundi mais meus estudos voltando eles para a área de técnicas de pré-processamento utilizadas neste contexto, me esbarrando na “pioneira” o MTCNN (Multitask Cascaded Convolutional Networks). Com isso, me familiarizei um pouco mais com seu uso e exatamente o que ela retornava e em quais formatos para poder começar a minha implementação de um pipeline completo de pré-processamento e treinamento “multitask”.

Com isso, durante esta Semana, realizei as seguintes atividades:

- Finalização do pipeline completo, com os seguintes aspectos todos implementados e validados:
 - Pré-processamento (detecção facial, crop, alinhamento facial e anotação dos “facial landmarks”
 - Treinamento multitask com o uso dos “facial landmarks”
- Com o código finalizado, realizei um teste para validar todo o processo.
- Foram utilizados os seguintes datasets: VggFace 2 112x112 (treino) e LFW (validação)
- Fiz alguns testes e percebi que não teria capacidade computacional para rodar o dataset inteiro (mais de 3 milhões de imagens), então decidi utilizar uma pequena porção para conseguir pelo menos validar o pipeline. Fiz um script para fazer um subset, gerando um resultado final de 90 identidade (de 8.631) e 62.558 imagens (de 3.310.000)
- Depois foi realizado o treinamento
- Com sua finalização, foi utilizado o script evaluate e as seguintes métricas foram obtidas:
 - Acurácia no dataset de validação = 92.17
 - Acurácia no LFW (Labeled Faces in the Wild) = 65.95%
 - AUC = 0.7176
 - TAR @ FAR = 0.001 = 2.87%
- Mesmo utilizando um subset extremamente reduzindo, o pipeline apresenta um caminho promissor

Documento com todo o processo realizado e resultados obtidos: [📄 TestePipeline](#)

Link para repositório com o código desenvolvido (acessar pasta Pipeline): [GitHub Pipeline](#)

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Como próximos passos, planejo estudar algumas melhorias para o pipeline que possam contribuir justamente para o quesito de aplicações de segurança, utilizando justamente a métrica TAR@FAR como parâmetro de avaliação. Além disso, procuro revisar algumas partes do processo de pré-processamento do pipeline que acredito que possam ficar ainda melhor.

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

Sofri “bastante” com a questão de força computacional que acabaram me atrasando bastante e prejudicando e interrompendo alguns treinamentos meus, além disso, me limitando a rodar em um conjunto extremamente pequeno de dados.

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: [Go! ▾](#)

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 5 de nov. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Durante as Semanas anteriores foram realizadas as seguintes atividades:

- Pesquisa de artigos e surveys a respeito do tema **Reconhecimento Facial**
- Construção de resumos e documentos com conhecimentos que envolvem a área
- Foco nos artigos que tratavam de técnicas de pré-processamento de imagens
- Testes iniciais com o MTCNN (Multitask Cascaded Convolutional Networks)
- Implementação de um pipeline completo de pré-processamento e treinamento

Já nesta Semana, foram realizadas as seguintes atividades:

- Ajustes no pipeline (especificamente na parte de pré-processamento) com o objetivo de melhorar os resultados obtidos. Após analisar as imagens, decidi fazer algumas alterações principalmente na parte do alinhamento facial. Mudanças feitas:
 - Alteração do uso da biblioteca do detector de face: uniface → facenet_pytorch
 - Algumas melhorias na função que realiza o alinhamento facial
 - Formato dos dados de detecção (facilitou a manipulação dos mesmos em etapas posteriores)
- Com isso, foi possível já obter melhores resultados que os apresentados anteriormente, utilizando os mesmos dados e configurações:
 - Acurácia no LFW (Labeled Faces in the Wild) = 65.95% → 69.10%
 - AUC = 0.7176 → 0.7496
 - TAR @ FAR = 0.001 = 2.87% → 4.30%
- Todo o processo acima pode ser encontrado no documento abaixo:
 - [TestePipeline2](#)
- Como a minha aplicação possui o foco em imagens, decidi aplicar uma nova técnica para atacar um ponto levantado anteriormente: Segurança. Com isso, decidi aplicar o Anti Spoofing (já citada anteriormente de forma breve em [Conhecimentos Base](#) e [Resumo Artigos](#))
- Início das implementações
 - [GitHub Pipeline](#)

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Para a próxima Semana planejo testar e validar o Anti-Spoofing implementado e, posteriormente, revisar o código e sua documentação.

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: Go! ▾

Teste Pipeline e Teste Pipeline 2

Resumo

Os documentos "Teste Pipeline" e "Teste Pipeline 2" descrevem duas iterações experimentais de um *pipeline* de reconhecimento facial baseado em aprendizado multitarefa (*multi-task learning*). Em ambos os experimentos, utilizou-se a arquitetura MobileNetV3-Large configurada com duas cabeças de saída: uma para a geração de *embeddings* faciais de 512 dimensões e outra para a predição de marcos faciais (*landmarks*). O conjunto de dados base para o treinamento foi o VGGFace2, que originalmente contém aproximadamente 3,3 milhões de imagens e 8.631 identidades. No entanto, devido a restrições computacionais e instabilidade de *hardware* relatadas, ambos os experimentos foram conduzidos utilizando uma subamostragem muito pequena, correspondendo a aproximadamente 1,5% do tamanho do *dataset* original, sendo a validação realizada no *benchmark* LFW.

No primeiro experimento, detalhado em "Resultados Treinamento Pipeline", a etapa de pré-processamento empregou o detector RetinaFace e uma função de perda combinada utilizando CosFace para classificação e SmoothL1 para os *landmarks*. Embora o processo tenha demonstrado viabilidade técnica com uma taxa de sucesso de pré-processamento superior a 98%, os resultados de generalização foram limitados pela escala reduzida dos dados. As métricas obtidas no melhor modelo (Época 11) apresentaram uma Acurácia de Verificação no LFW de 65,95%, uma AUC (*Area Under Curve*) de 0,7176 e uma taxa de aceitação verdadeira (TAR) com falsa aceitação (FAR) de 0,001 de 2,87%.

O documento "Resultados Treinamento Pipeline 2" apresenta otimizações significativas na etapa de pré-processamento, visando superar as limitações da primeira iteração. As principais alterações incluíram a substituição do detector RetinaFace pelo MTCNN, que reduziu o número de falhas de detecção, e o refinamento da função de

alinhamento facial para incluir controle explícito de interpolação e transformação automática de *landmarks*. Essas melhorias metodológicas resultaram em um ganho de desempenho consistente: o modelo alcançou uma Acurácia de Verificação de 69,10%, uma AUC de 0,7496 e uma TAR @ FAR=0,001 de 4,30%, representando uma evolução clara em relação ao *baseline* estabelecido no primeiro experimento.

APÊNDICE 6

Termo de Aceite de Entrega

Objetivo deste documento

Este documento faz parte do Processo da disciplina Residência em IA e tem como objetivo formalizar o aceite da entrega considerando o planejado e o realizado para o período.

Data da Reunião (“gate”) de aprovação: 13 de nov. de 2025

Participantes da Entrega [matriculados em Residência em IA]:

PEDRO RABELO MENDONÇA

Entrega: [descrever a ENTREGA: requisitos e produtos gerados: links para textos, códigos, vídeos etc.]

Durante as Semanas anteriores, foram realizadas as seguintes atividades:

- Pesquisa de artigos e surveys a respeito do tema **Reconhecimento Facial**
- Construção de resumos e documentos com conhecimentos que envolvem a área
- Foco nos artigos que tratavam de técnicas de pré-processamento de imagens
- Testes iniciais com o MTCNN (Multitask Cascaded Convolutional Networks)
- Implementação de um pipeline completo de pré-processamento e treinamento
- Melhorias alcançadas nos resultados obtidos

Agora, durante esta última Semana, comecei terminando as implementações e ajustes finais no meu Pipeline para implementar a terceira “cabeça” no treinamento dando assim suporte a detecção de fraudes, classificando uma imagem como real ou fake (uso de máscaras, adereços, fotos impressas, ...)

Realizei também mais treinamentos com a nova implementação. Foi feito um treinamento com o dataset específico para ataques de apresentação (CASIA-FASD) e, por fim, avaliações e inferências para verificar se o resultado esperado foi atingido com sucesso.

O objetivo principal era testar três cenários específicos para saber como o modelo se comportaria em cada um dele e verificar se ele teria sucesso em todos eles:

- Verificar se reconhece uma determinada pessoa (similaridade entre a imagem fornecida e a imagem presente no dataset)
- Verificar se diferencia duas pessoas distintas
- Verificar cenário de ataque em que uma pessoa tenta se passar pela outra

Em todos os três casos, o modelo não teve dificuldade em apresentar o resultado correto mesmo com uma baixa quantidade de dados de treinamento. Todos os resultados e parâmetros podem ser encontrados no seguinte documento: [Anti-Spoofing](#)

Planejamento: [descrever o que pretende fazer para realizar a próxima ENTREGA]

Observação: [caso precise fazer alguma observação, de qualquer “natureza”]

ACEITE DA ENTREGA:

CEDRIC LUIZ DE CARVALHO: Go! ▾

Anti-Spoofing

Resumo

O documento descreve a implementação e o treinamento de um sistema de reconhecimento facial integrado com capacidades de *anti-spoofing* (detecção de fraude), utilizando a arquitetura MobileNetV3-Large configurada para aprendizado multitarefa. Para viabilizar o experimento, foi criado um *dataset* híbrido totalizando 40.370 amostras, combinando 31.054 imagens do VGGFace2 (restrito a 90 identidades) para o reconhecimento facial e 9.316 amostras do CASIA-FASD para o treino de vivacidade. O modelo foi treinado ao longo de 30 épocas com funções de perda ponderadas para extração de embeddings, marcos faciais e *anti-spoofing*, alcançando uma acurácia de verificação no *benchmark* LFW de 65,80% e uma acurácia de teste no conjunto *anti-spoofing* (CASIA-FASD) de 84,60%.

Para validar a eficácia do sistema, foram realizados três testes de inferência em cenários distintos. No **Teste 1**, o modelo comparou duas imagens reais da mesma pessoa, validando corretamente a identidade e classificando ambas como "reais". O **Teste 2** avaliou duas imagens reais de pessoas diferentes, rejeitando a correspondência devido à baixa similaridade (0.057). O cenário mais crítico ocorreu no **Teste 3**, que simulou um ataque de apresentação onde uma foto impressa foi usada para tentar enganar o sistema; embora o modelo tenha detectado alta similaridade facial (0.73), ele rejeitou corretamente o acesso ao classificar a imagem de ataque como "FAKE" com um *score* de 0.82, demonstrando a funcionalidade da camada de segurança.

Conclusão

A conclusão do relatório avalia os resultados como satisfatórios, destacando que o modelo cumpriu com sucesso os objetivos de identificar indivíduos e distinguir entre faces reais e ataques de *spoofing*. Ressalto que, embora algumas métricas de performance (como a acurácia no LFW) tenham apresentado valores abaixo do estado da arte, isso é atribuído diretamente à escassez de dados de treinamento (apenas 90 identidades) imposta pelas limitações de recursos computacionais disponíveis, e não a uma falha na arquitetura do *pipeline* proposto.