

Revisão do Pré-projeto

Disciplina: Trabalho de Conclusão de Curso I – BCC

Caro, orientando,

segue abaixo o Termo de Compromisso, as DUAS revisões do seu pré-projeto contendo a avaliação do professor “avaliador” e professor “TCC1”, junto com as avaliações da defesa na banca de qualificação. É muito importante que revise com cuidado e discuta possíveis dúvidas decorrente das revisões com o seu professor orientador, e com o professor de TCC1. Sempre procure fazer todos os ajustes solicitados, até mesmo o menores detalhes, pois todos são importantes e irão refletir na sua nota nesta disciplina. Lembre de abrir localmente em um visualizador PDF para poder ver as anotações que foram feitas. E, aparecendo uma anotação feita por mim (prof. De TCC1) que inicie com “TF-...” (ex. “TF-ALÍNEA”) se refere a ajustes de formatação indicando que deve usar o estilo do Word correto do modelo do projeto.

Mas, caso o professor orientador julgue que algumas anotações das revisões não devam ser feitas, ou mesmo que sejam feitas de forma diferente a solicitada pelo revisor, anexe ao final do seu projeto a ficha “Projeto: Observações – Professor Orientador” disponível no material da disciplina, e justifique o motivo.

Lembrem que agora o limite de páginas do projeto é no máximo 12 (doze) páginas. E que a seção de “Revisão Bibliográfica” deve ser complementada (ampliada) usando subseções para cada assunto apresentado.

Atenciosamente,

UNIVERSIDADE REGIONAL DE BLUMENAU
CENTRO DE CIÊNCIAS EXATAS E NATURAIS
CURSO DE CIÊNCIA DA COMPUTAÇÃO – BACHARELADO

TERMO DE COMPROMISSO

I – IDENTIFICAÇÃO DO ALUNO	
Nome:	Jadiel dos Santos
CV Lattes:	http://lattes.cnpq.br/9010170549983604
E-mail:	jadielsantosjds@gmail.com
Telefone:	47 99211-3540
II – IDENTIFICAÇÃO DO TRABALHO	
Título provisório:	SISTEMA DE PRÉ SELEÇÃO DE FOTOS BASEADAS EM PADRÕES DE FACES
Orientador:	Aurélio Faustino Hoppe
Coorientador (se houver):	
Linha de Pesquisa:	<input type="checkbox"/> Tecnologias aplicadas à informática na educação <input checked="" type="checkbox"/> Tecnologias aplicadas ao desenvolvimento de sistemas
III – COMPROMISSO DE REALIZAÇÃO DO TCC	
Eu (aluno), Jadiel dos Santos comprometo-me a realizar o trabalho proposto no semestre 2023/1, de acordo com as normas e os prazos determinados pela FURB, conforme previsto na resolução nº.20/2016. Assinatura: NÃO É NECESSÁRIO – Encaminhar por mail ao orientador	
IV – COMPROMISSO DE ORIENTAÇÃO	
Eu (orientador), comprometo-me a orientar o trabalho proposto no semestre, de acordo com as normas e os prazos determinados pela FURB, conforme previsto na resolução nº.20/2016. Assinatura: NÃO É NECESSÁRIO – Encaminhar por mail ao professor de TCC I	

Blumenau, 17 de Agosto de 2022

CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC		
(X) PRÉ-PROJETO	() PROJETO	ANO/SEMESTRE: 2022/2

PROTÓTIPO DE VISÃO COMPUTACIONAL PARA SELEÇÃO DE FOTOS COM FACES

Jadiel dos Santos

Prof. Aurélio Faustino Hoppe - orientador

1 INTRODUÇÃO

Nós, seres humanos, somos excelentes em reconhecer instantaneamente uma grande variedade de objetos de forma automática, mesmo sem perceber, isso ocorre devido a capacidade de processamento do nosso cérebro em fazer operações complexas. Por isso, torna-se possível o reconhecimento de padrões (SANCHEZ-MORENO *et al.*, 2021). No entanto, ainda segundo os autores, para que computadores possam ser capazes de emular essa mesma atividade com a maestria dos seres humanos, muitos algoritmos e métodos foram propostos nas últimas décadas (SANCHEZ-MORENO *et al.*, 2021).

O reconhecimento facial é um conceito relativamente novo na computação, desenvolvido na década de 1960. Segundo He *et al.* (2019, p. 1), o reconhecimento facial, como um problema fundamental da área de visão computacional, tem sido bastante pesquisa nos últimos anos. O primeiro sistema semiautomatizado para o reconhecimento da face exigia que o seu administrador localizasse características nas fotografias (por exemplo, olhos, orelhas, nariz e boca) antes do sistema calcular distâncias e características para um ponto de referência comum, esse ponto de referência era depois comparado com os dados disponíveis.

Na literatura sobre reconhecimento facial, encontram-se diversas abordagens quanto a detecção, as quais, em sua maioria, se baseiam nas características do rosto, tais como a distância entre os olhos, comprimento do nariz, ossos laterais e cavidade orbital, entre outros. Já em relação as etapas realizadas para efetuar o reconhecimento, constam (i) a detecção das faces em uma imagem, (ii) extração de suas características, (iii) busca por imagens semelhantes em uma base de dados. Neste sentido, a grande problemática do reconhecimento facial encontra-se justamente em relação a detecção das faces na imagem. Oh *et al.* (2016) listam duas dimensões capazes de afetar a identificação de um indivíduo num sistema de reconhecimento facial: a quantidade de cabeças etiquetadas e o tipo de oclusão do rosto.

Atualmente, com o uso das Convolutional Neural Network (CNNs), a acurácia para reconhecimento em ambientes controlados aumentou significativamente. Assim, Yang *et al.* (2018, p. 1) afirmam que este método, juntamente com técnicas de processamento de imagem podem efetivamente melhorar a capacidade de uma rede de detecção de rosto em lidar com oclusões graves. Diante deste contexto, o reconhecimento facial está atraindo atenção de vários mercados, sendo aplicado para funções de marketing, segurança e seleção ou busca de conteúdos por fotos.

Em face do exposto, este trabalho visa utilizar redes neurais convolucionais para selecionar fotografias ao qual um determinado indivíduo aparece. Desta forma, os profissionais que fotografam eventos como formaturas poderão informar uma face como entrada e o protótipo efetuará a pré-seleção das fotos.

1.1 OBJETIVOS

O objetivo deste trabalho é desenvolver um protótipo de reconhecimento facial que, a partir de uma imagem de entrada, retorne todas as fotos aos quais encontram-se um determinado indivíduo.

Os objetivos específicos são:

- investigar as técnicas mais utilizadas na detecção e reconhecimento de faces;
- analisar a qualidade mínima das imagens para realizar a detecção facial;
- analisar e comparar a precisão dos algoritmos utilizados na seleção e agrupamento as imagens por indivíduo.

2 TRABALHOS CORRELATOS

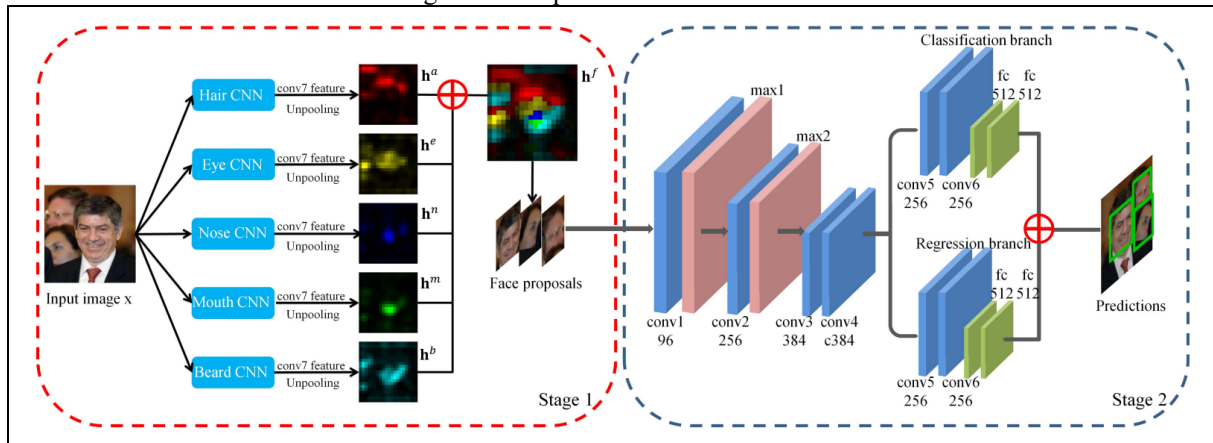
Nesta seção serão apresentados trabalhos com características semelhantes aos principais objetivos do estudo proposto. A seção 2.1 descreve o processo de detecção de faces por meio das regiões e atributos faciais: cabelo, olho, nariz, boca e barba (YANG *et al.*, 2018). A seção 2.2 discorre sobre o método de reconhecimento de faces implementado através de uma rede de Face Deformável (HE *et al.*, 2019). Por fim, a seção 2.3 apresenta o método para reconhecimento facial em vídeos e fotos em tempo real (SANCHEZ-MORENO *et al.*, 2021).

2.1 FACENESS-NET: FACE DETECTION THROUGH DEEP FACIAL PART RESPONSES

Yang *et al.* (2018) propuseram um método para melhorar o processo de detecção facial. Segundo os autores, embora as *bounding boxes* de face sejam utilizadas como padrão de entrada em detectores de faces, a utilidade desses atributos ainda é pouco explorada. Neste sentido, eles apresentam o conceito de que a supervisão de atributos faciais pode aumentar a capacidade de uma *rede neural* em lidar com oclusões e posições diversas (YANG *et al.*, 2018).

Segundo Yang *et al.* (2018), utilizando-se do conceito de supervisão dos atributos faciais, criou-se a Faceness-Net que possui dois estágios: (i) a *convolutional neural network* (CNN) recebe uma imagem como *input* e gera um *partness map*¹, indicando 5 partes da face (cabelos, olhos, nariz, boca e barba). A partir disso, se estabelece o ranqueamento de algumas janelas candidatas utilizando pontos faciais; (ii) as janelas candidatas são refinadas objetivando otimizar as classificações de face e as *bounding box regression* (YANG *et al.*, 2018). A Figura 1 demonstra os principais elementos da arquitetura Faceness-Net.

Figura 1 – Arquitetura da Faceness-Net



Fonte: Yang *et al.* (2018).

Yang *et al.* (2018) também apontam que o primeiro estágio da linha base da Faceness-Net consiste de múltiplas redes de reconhecimento de atributos para gerar mapas de resposta de diferentes partes da imagem. Para isso são utilizadas 5 redes, uma para cada componente da face, as quais compartilham da mesma estrutura com 7 camadas de convolução e 2 de *max-pooling*. Ao fim da sétima convolução é gerado o *response map*, no qual aplica-se a normalização L2 e média ao longo dos canais, indicando a localização do componente facial. Posteriormente, realiza-se o *unpooling* para que seja obtido o *partness map* que será do mesmo tamanho da imagem inicial. Dessa forma, soma-se os *partness map* para que sejam obtidas janelas da *face proposals* (YANG *et al.*, 2018).

Segundo Yang *et al.* (2018), a segunda etapa do Faceness-Net refina a janela de candidatos gerada a partir do primeiro estágio usando uma CNN multitarefa, onde a classificação de face e a *bounding box regression* são otimizadas em conjunto. Esta CNN multitarefa é pré-treinada nas camadas convolucionais de 1 a 4 do primeiro estágio. Na quarta camada, a CNN é dividida em 2 ramos, os quais consistem em 2 camadas convolucionais. Um destes ramos irá executar a classificação de face, e o outro fará a *bounding box regression* (YANG *et al.*, 2018).

Nos experimentos realizados, o método obteve bons resultados nos *datasets Annotated Faces in the Wild* (AFW) (ZHU; RAMANAN, 2012), *Pattern Analysis, Statistical Modelling and Computational Learning* (PASCAL) (NADA *et al.*, 2018), e WIDER FACE (YANG *et al.*, 2016). Sua acurácia no subconjunto de nível fácil foi de, respectivamente, 98,05%, 92,11% e 71,7% (YANG *et al.*, 2018).

De acordo com Yang *et al.* (2018), estas redes de reconhecimento de atributos podem compartilhar camadas de convolução iniciais, assim reduzindo em 83% os parâmetros, bem como as computações, além de melhorar a robustez para a detecção facial. Por fim, a eficácia do método também pode ser atingida a partir de atributos de imagens de faces não cortadas, sem supervisão explícita das partes.” (YANG *et al.*, 2018, p. 12, tradução nossa). Além disso, eles também destacam que obtiveram sucesso ao introduzir o conceito de *faceness score*, o qual pode ser usado para ranquear janelas candidatas em qualquer técnica de proposta de região, podendo gerar conjuntos de propostas faciais com alta taxa de recall (YANG *et al.*, 2018).

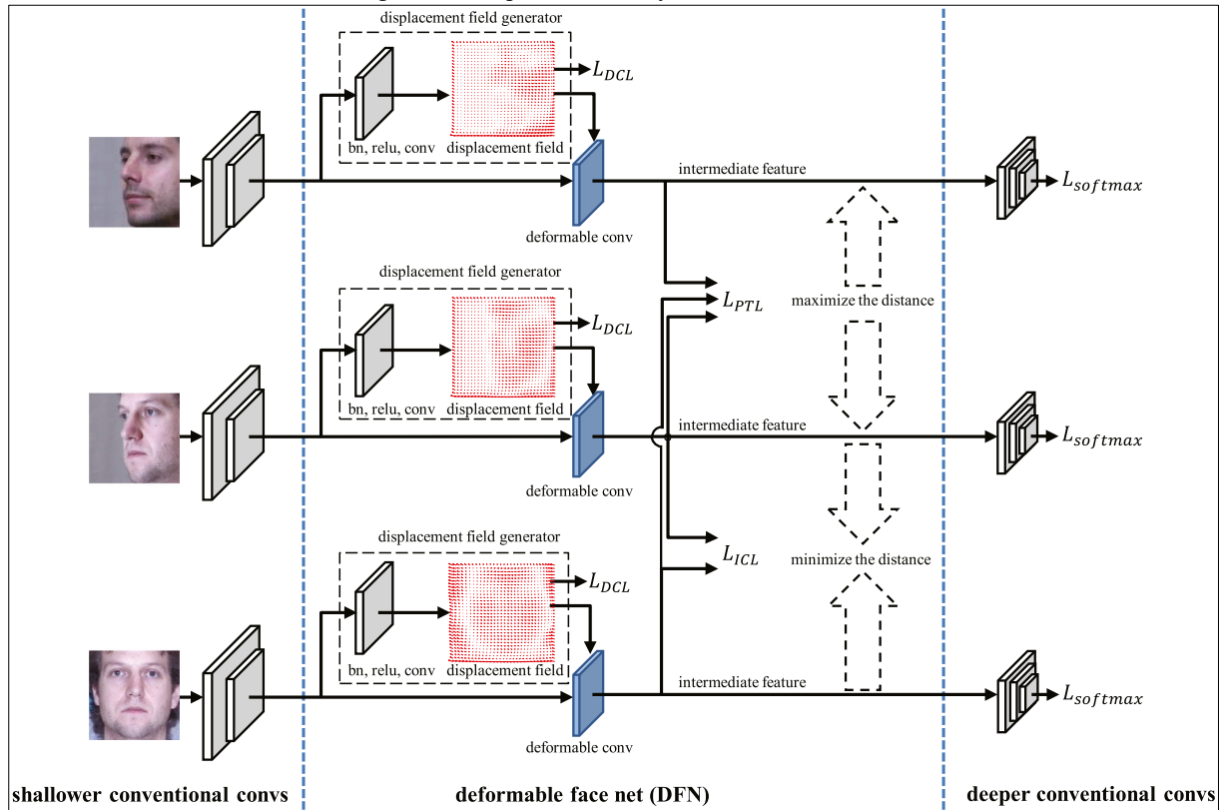
¹ Partness map: São os mapas de resposta que destacam 5 partes diferentes da face (YANG *et al.*, 2018, p. 1-2, tradução nossa).

2.2 DEFORMABLE FACE NET FOR POSE INVARIANT FACE RECOGNITION

Segundo He *et al.* (2019), a acurácia do reconhecimento facial tem aumentado rapidamente com o uso de *Convolutional Neural Network* (CNN). Entretanto, os autores ressaltam que sob circunstâncias não controladas, como poses, iluminação, resolução, oclusão e expressão, a performance do sistema de reconhecimento facial é afetada significativamente. Por conta disso, He *et al.* (2019) propuseram um método de alinhamento em nível de *feature* para lidar com variações de pose no reconhecimento de rosto.

He *et al.* (2019) ressaltam que o método inicialmente aplica uma *Deformable Face Net* (DFN), a qual busca aprender simultaneamente o alinhamento em nível de *feature* e a extração de *features* para reconhecimento de rosto por meio de convoluções deformáveis com um campo de deslocamento espacial. Dessa maneira, estes campos permitem que a DFN seja capaz de alinhar as *features*, independente de variações de pose. A Figura 2 apresenta a arquitetura da DFN.

Figura 2 – Arquitetura da *Deformable Face Net*



Fonte: He *et al.* (2019).

A partir da Figura 2, pode-se observar que a DFN é composta de uma camada de convolução inicial com função de ativação *ReLU* e com a função de perda *Displacement Consistency Loss* (DCL) para que seja gerado um *displacement field*. Em seguida, passa por uma *deformable convolution*, a qual gera uma *feature* intermediária otimizada pelas funções de perda *Identity Consistency Loss* (ICL) e *Pose-Triplet Loss* (PTL). Por fim, as camadas mais profundas de convolução utilizam a função de perda *Softmax*.

Com o propósito de que os campos de deslocamento espacial sejam consistentes em sua orientação e amplitude, a DCL é utilizada para ~~que~~ obter um alinhamento melhor para reconhecer a face (HE *et al.*, 2019). Neste sentido, as funções ICL e PTL, servem para minimizar a variação de *features* intraclasse causada por poses diferentes e maximizar a distância de *recursos* entre classes sob poses semelhantes. Dessa forma, He *et al.* (2019) afirmam que a DFN supera os métodos de última geração nos seguintes *datasets* de testes: MegaFace (KEMELMACHER-SHLIZERMAN *et al.*, 2019), MultiPIE (GROSS *et al.*, 2008) e CFP (SENGUPTA *et al.*, 2016). No MegaFace, o método obteve 82,11% de precisão. No MultiPIE, obteve-se nas poses de $\pm 90^\circ$ uma acurácia de 85,66%. Por fim, o *dataset* CFP, obteve-se uma acurácia de 94,01%. Com isso, He *et al.* (2019) afirmam que o método aumenta significativamente o reconhecimento de rosto em todas as poses.

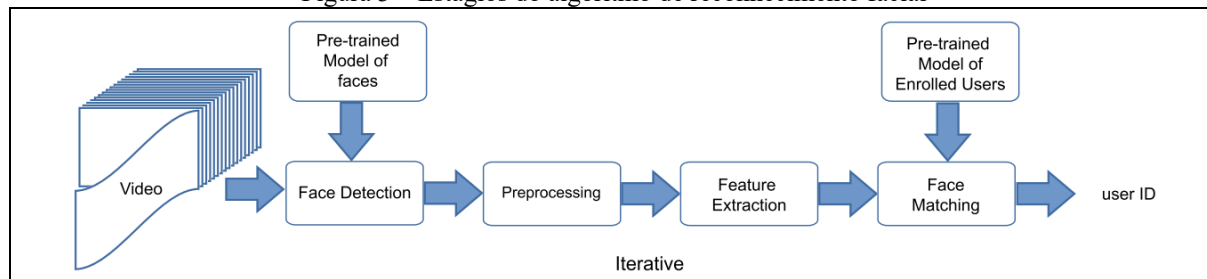
2.3 EFFICIENT FACE RECOGNITION SYSTEM FOR OPERATING IN UNCONSTRAINED ENVIRONMENTS

Sanchez-Moreno *et al.* (2021) desenvolveram um ~~de~~-sistema reconhecimento facial para ambientes irrestritos. Eles utilizaram o YOLO-Face modificado com um esquema de classificação que permite aumentar a acurácia da detecção com redução do tempo computacional. Os autores também propuseram uma arquitetura que

combina uma CNN com um algoritmo de aprendizagem supervisionada, como por exemplo, o *Support Vector Machine* (SVM).

Segundo Sanchez-Moreno *et al.* (2021), subdividiu-se o algoritmo em 4 estágios: No primeiro estágio, realiza-se a detecção da face por meio de um *input* de *frame* de vídeo ou imagem. Para isso, eles utilizaram o YOLO-Face, ao qual detecta as faces em tempo real. Desse modo, os autores resolveram o problema de velocidade. A Figura 3 apresenta os estágios do algoritmo de reconhecimento facial.

Figura 3 – Estágios do algoritmo de reconhecimento facial



Fonte: Sanchez-Moreno *et al.* (2021).

No segundo estágio, realiza-se o redimensionamento, normalização L2 e conversão da imagem para escala de cinza, tendo como intuito a preparação para aplicar o *machine learning model* (SANCHEZ-MORENO *et al.* 2021, p. 4-7). No terceiro estágio ocorre a extração das *features* através dos métodos de classificação como Support Vector Machine (SVM), Random Forest (RF) e K-Nearest Neighbors (KNN). Por fim, no último estágio, as *features* faciais são comparadas com as *features* conhecidas de usuários cadastrados para identificação pessoal.

Para fins de avaliação do reconhecimento facial, Sanchez-Moreno *et al.* (2021) utilizaram os seguintes *datasets*: *Labeled Faces in the Wild* (LFW) (HUANG *et al.*, 2007) e *YouTube Faces* (YTF) (WOLF; HASSNER; MAOZ, 2011). De acordo com os autores, no LFW, o método obteve uma acurácia de 99,1%. Já no YTF, a acurácia foi de 94,7%.

A partir desses resultados, Sanchez-Moreno *et al.* (2021) concluem que o sistema proposto melhora a precisão no banco de dados LFW, obtendo pontuação semelhante ou superior a diferentes estruturas de CNN utilizadas em reconhecimentos faciais. Assim, no que diz respeito ao reconhecimento facial, os autores resumiram que “[...] com os avanços dos algoritmos de deep learning aliados aos avanços tecnológicos, é possível desenvolver sistemas de reconhecimento facial em tempo real capazes de operar em ambientes de alta complexidade.” (SANCHEZ-MORENO *et al.*, 2021).

3 PROPOSTA

Esta seção apresentada a justificativa para o desenvolvimento desse trabalho, os principais requisitos e a metodologia de desenvolvimento que será utilizada. Também são relacionados os assuntos e as fontes bibliográficas que irão fundamentar o estudo proposto.

3.1 JUSTIFICATIVA

No Quadro 1 é apresentado o comparativo entre os trabalhos correlatos. As linhas representam as características e as colunas os trabalhos.

Quadro 1 - Comparativo dos trabalhos correlatos

Características \ Trabalhos correlatos	Yang <i>et al.</i> (2018)	He <i>et al.</i> (2019)	Sanchez-Moreno <i>et al.</i> (2021)
Detecção de faces	Sim	Sim	Sim
Reconhecimento de faces	Não	Não	Sim
Uso das regiões faciais para aprendizado	Sim	Não	Não
Lida com ambientes irrestritos (oclusões, iluminação, dimensionamento e poses)	Sim	Sim	Sim
Arquitetura de rede neural	Faceness-Net	DFN	YOLO-faces
Datasets utilizados	AFW, PASCAL, Fddb e WIDER FACE	MegaFace, MultiPIE e CFP	LFW e YTF
Taxa de acurácia para os datasets	98,05%, 92,11% e 71,7%	82,11%, 85,66% e 94,01%.	99,1% e 94,7%

Fonte: elaborado pelo autor.

A partir do Quadro 1, pode-se observar que todos os estudos utilizam de **redes neurais convolucionais (CNN)** como método de aprendizado de máquina. No que diz respeito a detecção de faces, todos os estudos correlatos implementam diferentes métodos. Yang *et al.* (2018) utilizaram atributos faciais para classificar e detectar faces. He *et al.* (2019) optaram pela rede *Deformable Face Net* (DFN), a qual simultaneamente alinha e extrai as *features*. Já Sanchez-Moreno *et al.* (2021) utilizaram o método YOLO-faces, que calcula a probabilidade em cada região, destacando-as através de *bounding boxes*.

No que diz respeito ao reconhecimento facial, nota-se que apenas Sanchez-Moreno *et al.* (2021), implementam tal funcionalidade. Já Yang *et al.* (2018), com seu método *Faceness-Net*, utilizam regiões faciais, ou seja, cabelo, olhos, nariz, boca e barba para realizar a detecção facial. Já em relação aos ambientes irrestritos, isto é, levar em conta as oclusões, poses variadas, dimensões da imagem e iluminação, todos os estudos implementaram solução as quais conseguem obter bons resultados nestas condições adversas. Além disso pode-se observar que a acurácia obtida pelo método de Yang *et al.* (2018) nos *datasets* AFW, PASCAL, FDDB e WIDER FACE foi de 98,05%, 92,11% e 71,7%. Já as acurácias atingidas pelo método de He *et al.* (2019) nos *datasets* MegaFace, MultiPIE e CFP foram de 82,11%, 85,66% e 94,01%. E, Sanchez-Moreno *et al.* (2021) obtiveram 99,1% e 94,7% nos *datasets* LFW e YTF.

A partir deste contexto, este trabalho implementará um protótipo de detecção e reconhecimento facial em ambientes irrestritos para que, dado um conjunto de imagens, seja feita a pré-seleção automática para cada indivíduo informado. O protótipo a ser desenvolvido poderá ser utilizado por fotógrafos, aos quais após a realização de eventos como formaturas, casamentos e festas em geral, precisam realizar o processo de verificação de todas as imagens retiradas para encontrar e agrupar as fotos de acordo com quem aparece nas imagens. Ressalta-se que tal processo demanda muito tempo para ser feito manualmente, pois em um evento são tiradas milhares de fotos. Além disso, também se destaca a necessidade de estudos para definir qual é a qualidade mínima para o processamento e reconhecimento ocorrer com melhor precisão.

No cenário de fotografias profissionais, há alguns anos, o processo de envio das fotos tiradas pelo fotógrafo para que o cliente fizesse a seleção das mesmas era feito através de CDs ou DVDs, sendo posteriormente substituídos por *pendrives*. Deste modo, todos os fotógrafos que querem oferecer um serviço profissional e de qualidade para seu cliente, buscam formas de melhorar todos os processos que envolvem seu trabalho como um todo, e a tecnologia, técnicas de processamentos e aprendizado e máquina podem ajudá-los. Assim, este trabalho tem como essência a otimização do processo de seleção das fotografias, ao qual seus clientes apareçam nas imagens. Para isso, serão investigados e utilizados algoritmos de reconhecimento facial e redes neurais convolucionais que se adequem melhor a ambientes irrestritos.

3.2 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

Esta seção mostrará os Requisitos Funcionais (RF) e Requisitos Não Funcionais (RNF) para a implementação do protótipo de seleção de fotos a partir de faces. Os requisitos do trabalho estão divididos em duas partes: **requisitos da aplicação web** e requisitos do módulo de reconhecimento.

A aplicação web deverá:

- a) manter o cadastro de usuários (RF);
- b) permitir o login de usuários já cadastrados (RF);
- c) permitir ao usuário a criação de uma nova galeria de fotos (RF);
- d) permitir ao usuário visualizar as fotos pré-selecionadas (RF);
- e) utilizar o ambiente de desenvolvimento Visual Studio Code (RNF);
- f) desenvolver a aplicação web utilizando NodeJS no backend e React no frontend (RNF);
- g) utilizar o banco de dados PostgreSQL para persistir os dados (RNF).

O módulo de reconhecimento deverá:

- a) utilizar técnicas de processamento de imagens para melhorar a qualidade das imagens obtidas (filtragem de ruído e normalização da iluminação) (RF);
- b) realizar a normalização e a extração de características para estabelecer a qualidade mínima necessária da imagem para realizar a detecção e o reconhecimento facial (RF);
- c) utilizar algoritmos de aprendizado de máquina para efetuar a detecção das faces e consecutivamente a seleção das fotos a partir da face dos indivíduos (RF);
- d) utilizar métricas de similaridade para comparar as faces de entrada com os presentes nas fotos (RF);
- e) utilizar métricas de classificação: PCA, LPA, ICA, SVM, RF, KNN ou identificar outro classificador que produza melhor resultado (RF);
- f) ser desenvolvida na linguagem de programação Python (RNF);
- g) utilizar as bibliotecas OpenCV, Keras e scikit-learn (RNF).

3.3 METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

- levantamento bibliográfico: pesquisar sobre técnicas e novas tecnologias para reconhecimento facial de pessoal, processamento de imagens para melhoria **no** reconhecimento, aprendizado de máquina e trabalhos correlatos;
- elicitação dos requisitos da aplicação web: baseando-se nas informações da etapa anterior, reavaliar os requisitos propostos para a aplicação, tanto funcionais como não funcionais;
- especificação: utilizar a ferramenta de diagramação Enterprise Architect (EA) para elaborar os diagramas de caso de uso e de atividades de acordo com a Unified Modeling Language (UML);
- implementação: a partir do item (c) implementar a aplicação web utilizando o NodeJS no *backend* e React no *frontend*;
- testes da aplicação web: elaborar testes para validar a usabilidade da aplicação web;
- coleta e montagem da base de dados: coletar fotos/vídeos ou utilizar bases de dados com faces obtidas/cedidas junto a projetos já existentes na comunidade acadêmica;
- preparação das imagens: realizar o realce e melhoramento das imagens utilizando o OpenCV. Posteriormente subdividir as imagens em dois grupos: imagens de treinamento e de validação.
- definição da arquitetura da rede neural convolucional: definir a arquitetura da rede neural convolucional mais aderente ao processo de reconhecimento facial;
- implementação do reconhecimento facial: a partir do item (h) realizar a implementação do módulo de reconhecimento facial utilizando OpenCV para o processamento das imagens e a biblioteca Keras para aprendizado de máquina;
- testes do reconhecimento facial: paralelamente à implementação do item (i), realizar os testes do módulo de reconhecimento facial, com base em um banco de imagens de familiares e amigos.

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro 2 – Cronograma de atividades a serem realizadas

etapas / quinzenas	2023											
	jan.		fev.		mar.		abr.		maio		jun.	
	1	2	1	2	1	2	1	2	1	2	1	2
levantamento bibliográfico												
elicitação de requisitos da aplicação web												
especificação												
implementação												
testes da aplicação web												
coleta e montagem da base de dados												
preparação das imagens												
definição da arquitetura da rede neural convolucional												
implementação do reconhecimento facial												
testes do reconhecimento facial												

Fonte: elaborado pelo autor.

4 REVISÃO BIBLIOGRÁFICA

Nesta seção encontra-se uma breve descrição dos assuntos relacionados a este estudo: **Reconhecimento facial, visão computacional e Rede Neural Convolucional.**

O reconhecimento facial é uma técnica de biometria baseada nos traços do rosto das pessoas. Não é segredo que os seres humanos têm uma facilidade para reconhecer os rostos das pessoas com precisão. Mesmo que duas pessoas sejam parecidas, nós somos capazes de distingui-las. Essa técnica pode ser transmitida para os computadores, mas, para isso, precisa-se definir traços únicos que devem ser mapeados em códigos binários. (PISA, 2022).

Segundo Sanchez-Moreno *et al.* (2021), os sistemas de reconhecimento facial, devem primeiro detectar uma face independente do meio em que se encontra, para que finalmente possa comparar as *features* extraídas com as faces cadastradas previamente, assim fazendo a correspondência de face. Ainda de acordo com os autores, uma das aplicações para o reconhecimento facial é seu uso em sistemas de segurança em tempo real. Além disso, Sanchez-Moreno *et al.* (2021) também afirmam que os algoritmos ou métodos que utilizam inteligência artificial, mais especificamente redes neurais, emulam a capacidade do ser humano de reconhecer uma imensa variedade de objetos, animais, faces, etc.

Segundo Rosa (2018), a Rede Neural Convolucional (do inglês *Convolutional Neural Network* - CNN) é uma classe da rede neural artificial que começou a ser aplicada no processamento e análise de imagens digitais. A CNN é uma variante da rede *perceptrons* de múltiplas camadas, e foi inspirada no processo biológico de

processamentos de dados visuais. As células responsáveis pelo processamento dos dados visuais são sensíveis a pequenas sub-regiões do campo visual, também chamado de campo receptivo. Essas células são adequadas para explorar a correlação espacial local presente em imagens naturais, visto que elas atuam como filtros locais sobre o campo de visão. O campo de visão é todo coberto através dos campos receptivos de diferentes neurônios que se sobrepõem parcialmente. Imagens coloridas em treinamentos podem resultar numa melhora significativa na interpretação dos dados, pois a CNN tira maior proveito nestas imagens do que em imagens em escala de cinza. Isto ocorre porque assim como nos humanos, a cor ajuda na análise e interpretação de objetos, facilitando a memorização e reconhecimento.

REFERÊNCIAS

- GROSS, Ralph *et al.*, Multi-PIE. In: IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE & GESTURE RECOGNITION, 8., 2008, Amsterdam. **Proceedings...** Amsterdam, 2008. p. 1-8.
- HE, Mingjie *et al.*, Deformable Face Net: Learning Pose Invariant Feature with Pose Aware Feature Alignment for Face Recognition. In: IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE & GESTURE RECOGNITION, 14., 2019, Lille. **Proceedings...** Lille: Lille grand Palais, 2019. p. 1-8.
- HUANG, Gary B. *et al.* **Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments**. 2007. 11 f. Technical Report 07-49 - University of Massachusetts, Amherst.
- KEMELMACHER-SHLIZERMAN, Ira *et al.*, **The MegaFace Benchmark: 1 Million Faces for Recognition at Scale**. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 29., 2016, Las Vegas. **Proceedings...** Las Vegas: Caesars Palace, 2016. p. 4873-4882.
- NADA, Hajime *et al.*, **Pushing the Limits of Unconstrained Face Detection: a Challenge Dataset and Baseline Results**. 2018. Disponível em: <<https://arxiv.org/abs/1804.10275>>. Acesso em: 27 set. 2022
- PISA, Pedro. **Como funciona o reconhecimento facial**. Disponível em: <<http://www.techtudo.com.br/artigos/noticia/2012/04/como-funciona-o-reconhecimento-facial.html>>. Acessado em: 13 set. 2022
- ROSA, Renan P. **Método de classificação de pragas por meio de rede neural convolucional profunda**. 2018. 101 f. Dissertação (Mestrado em Computação Aplicada) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2018.
- SANCHEZ-MORENO, Alejandra S. *et al.* 2021. Efficient Face Recognition System for Operating in Unconstrained Environments. **Journal of Imaging**, v. 7, n. 9, p. 161, ago. 2021.
- SENGUPTA, Soumyadip *et al.*, Frontal to profile face verification in the wild. In: IEEE WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION, 16., 2016, Lake Placid. **Proceedings...** Lake Placid: Crowne Plaza, 2016. p. 1-9.
- WOLF, Lior; HASSNER, Tal; MAOZ, Itay. Face Recognition in Unconstrained Videos with Matched Background Similarity. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 24., 2011, Colorado Springs. **Proceedings...** Colorado Springs: Crowne Plaza, 2011. p. 529-534.
- YANG, Shuo *et al.* Faceness-Net: Face Detection through Deep Facial Part Responses. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 40, n. 8, p. 1845-1859, ago. 2018.
- YANG, Shuo *et al.*, **Wider Face: A Face Detection Benchmark**. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 29., 2016, Las Vegas. **Proceedings...** Las Vegas: Caesars Palace, 2016. p. 5525-5533.
- ZHU, Xiangxin; RAMANAN, Deva. Face detection, Pose Estimation, and Landmark Localization in the Wild. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 25., 2012, Providence. **Proceedings...** Providence: Rhode Island Convention Center, 2012. p. 2879-2886.

FORMULÁRIO DE AVALIAÇÃO BCC – PROFESSOR AVALIADOR – PRÉ-PROJETO

Avaliador(a): Andreza Sartori

Atenção: quando o avaliador marcar algum item como atende parcialmente ou não atende, deve obrigatoriamente indicar os motivos no texto, para que o aluno saiba o porquê da avaliação.

ASPECTOS AVALIADOS		Atende	atende parcialmente	não atende
ASPECTOS TÉCNICOS	1. INTRODUÇÃO O tema de pesquisa está devidamente contextualizado/delimitado?			
	O problema está claramente formulado?			
	2. OBJETIVOS O objetivo principal está claramente definido e é passível de ser alcançado?			
	Os objetivos específicos são coerentes com o objetivo principal?			
	3. TRABALHOS CORRELATOS São apresentados trabalhos correlatos, bem como descritas as principais funcionalidades e os pontos fortes e fracos?			
	4. JUSTIFICATIVA Foi apresentado e discutido um quadro relacionando os trabalhos correlatos e suas principais funcionalidades com a proposta apresentada?			
	São apresentados argumentos científicos, técnicos ou metodológicos que justificam a proposta?			
	São apresentadas as contribuições teóricas, práticas ou sociais que justificam a proposta?			
	5. REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO Os requisitos funcionais e não funcionais foram claramente descritos?			
	6. METODOLOGIA Foram relacionadas todas as etapas necessárias para o desenvolvimento do TCC?			
	Os métodos, recursos e o cronograma estão devidamente apresentados e são compatíveis com a metodologia proposta?			
	7. REVISÃO BIBLIOGRÁFICA Os assuntos apresentados são suficientes e têm relação com o tema do TCC?			
	As referências contemplam adequadamente os assuntos abordados (são indicadas obras atualizadas e as mais importantes da área)?			
ASPECTOS METODOLÓGICOS	8. LINGUAGEM USADA (redação) O texto completo é coerente e redigido corretamente em língua portuguesa, usando linguagem formal/científica?			
	A exposição do assunto é ordenada (as ideias estão bem encadeadas e a linguagem utilizada é clara)?			

CURSO DE CIÊNCIA DA COMPUTAÇÃO – TCC	
(X) PRÉ-PROJETO () PROJETO	ANO/SEMESTRE: 2022/2

PROTÓTIPO DE VISÃO COMPUTACIONAL PARA SELEÇÃO DE FOTOS COM FACES

Jadiel dos Santos

Prof. Aurélio Faustino Hoppe - orientador

1 INTRODUÇÃO

Nós, seres humanos, somos excelentes em reconhecer instantaneamente uma grande variedade de objetos de forma automática, mesmo sem perceber, isso ocorre devido a capacidade de processamento do nosso cérebro em fazer operações complexas. Por isso, torna-se possível o reconhecimento de padrões (SANCHEZ-MORENO *et al.*, 2021). No entanto, ainda segundo os autores, para que computadores possam ser capazes de emular essa mesma atividade com a maestria dos seres humanos, muitos algoritmos e métodos foram propostos nas últimas décadas (SANCHEZ-MORENO *et al.*, 2021).

O reconhecimento facial é um conceito relativamente novo na computação, desenvolvido na década de 1960. Segundo He *et al.* (2019, p. 1), o reconhecimento facial, como um problema fundamental da área de visão computacional, tem sido bastante pesquisado nos últimos anos. O primeiro sistema semiautomatizado para o reconhecimento da face exigia que o seu administrador localizasse características nas fotografias (por exemplo, olhos, orelhas, nariz e boca) antes do sistema calcular distâncias e características para um ponto de referência comum, esse ponto de referência era depois comparado com os dados disponíveis.

Na literatura sobre reconhecimento facial, encontram-se diversas abordagens quanto a detecção, as quais, em sua maioria, se baseiam nas características do rosto, tais como a distância entre os olhos, comprimento do nariz, ossos laterais e cavidade orbital, entre outros. Já em relação as etapas realizadas para efetuar o reconhecimento, constam (i) a detecção das faces em uma imagem, (ii) extração de suas características, (iii) busca por imagens semelhantes em uma base de dados. Neste sentido, a grande problemática do reconhecimento facial encontra-se justamente em relação a detecção das faces na imagem. Oh *et al.* (2016) listam duas dimensões capazes de afetar a identificação de um indivíduo num sistema de reconhecimento facial: a quantidade de cabeças etiquetadas e o tipo de oclusão do rosto.

Comentado [DSdR1]: Citação não encontrada nas referências.

Atualmente, com o uso das Convolutional Neural Network (CNNs), a acurácia para reconhecimento em ambientes controlados aumentou significativamente. Assim, Yang *et al.* (2018, p. 1) afirmam que este método, juntamente com técnicas de processamento de imagem podem efetivamente melhorar a capacidade de uma rede de detecção de rosto em lidar com oclusões graves. Diante deste contexto, o reconhecimento facial está atraindo atenção de vários mercados, sendo aplicado para funções de marketing, segurança e seleção ou busca de conteúdos por fotos.

Em face do exposto, este trabalho visa utilizar redes neurais convolucionais para selecionar fotografias ao qual um determinado indivíduo aparece. Desta forma, os profissionais que fotografam eventos como formaturas poderão informar uma face como entrada e o protótipo efetuará a pré-seleção das fotos.

1.1 OBJETIVOS

O objetivo deste trabalho é desenvolver um protótipo de reconhecimento facial que, a partir de uma imagem de entrada, retorne todas as fotos aos quais encontram-se um determinado indivíduo.

Os objetivos específicos são:

- investigar as técnicas mais utilizadas na detecção e reconhecimento de faces;
- analisar a qualidade mínima das imagens para realizar a detecção facial;
- analisar e comparar a precisão dos algoritmos utilizados na seleção e agrupamento as imagens por indivíduo.

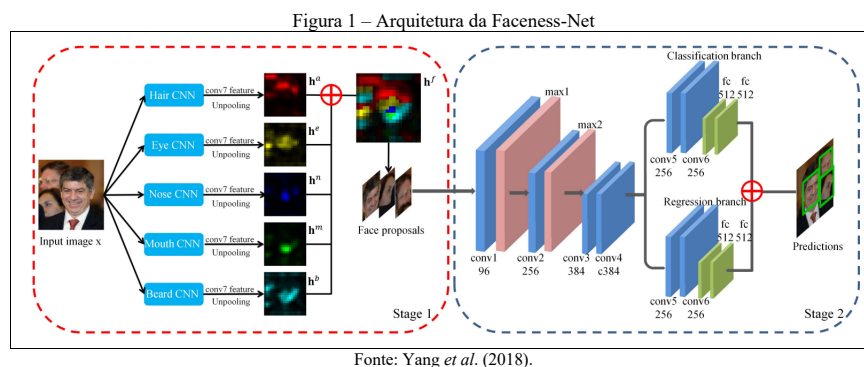
2 TRABALHOS CORRELATOS

Nesta seção serão apresentados trabalhos com características semelhantes aos principais objetivos do estudo proposto. A seção 2.1 descreve o processo de detecção de faces por meio das regiões e atributos faciais: cabelo, olho, nariz, boca e barba (YANG *et al.*, 2018). A seção 2.2 discorre sobre o método de reconhecimento de faces implementado através de uma rede de Face Deformável (HE *et al.*, 2019). Por fim, a seção 2.3 apresenta o método para reconhecimento facial em vídeos e fotos em tempo real (SANCHEZ-MORENO *et al.*, 2021).

2.1 FACENESS-NET: FACE DETECTION THROUGH DEEP FACIAL PART RESPONSES

Yang *et al.* (2018) propuseram um método para melhorar o processo de detecção facial. Segundo os autores, embora as *bounding boxes* de face sejam utilizadas como padrão de entrada em detectores de faces, a utilidade desses atributos ainda é pouco explorada. Neste sentido, eles apresentam o conceito de que a supervisão de atributos faciais pode aumentar a capacidade de uma rede neural em lidar com oclusões e posições diversas (YANG *et al.*, 2018).

Segundo Yang *et al.* (2018), utilizando-se do conceito de supervisão dos atributos faciais, criou-se a Faceness-Net que possui dois estágios: (i) a *convolutional neural network* (CNN) recebe uma imagem como *input* e gera um *partness map*¹, indicando 5 partes da face (cabelos, olhos, nariz, boca e barba). A partir disso, se estabelece o ranqueamento de algumas janelas candidatas utilizando pontos faciais; (ii) as janelas candidatas são refinadas objetivando otimizar as classificações de face e as *bounding box regression* (YANG *et al.*, 2018). A Figura 1 demonstra os principais elementos da arquitetura Faceness-Net.



Yang *et al.* (2018) também apontam que o primeiro estágio da linha base da Faceness-Net consiste em múltiplas redes de reconhecimento de atributos para gerar mapas de resposta de diferentes partes da imagem. Para isso são utilizadas 5 redes, uma para cada componente da face, as quais compartilham da mesma estrutura com 7 camadas de convolução e 2 de *max-pooling*. Ao fim da sétima convolução é gerado o *response map*, no qual aplica-se a normalização L2 e média ao longo dos canais, indicando a localização do componente facial. Posteriormente, realiza-se o *unpooling* para que seja obtido o *partness map* que será do mesmo tamanho da imagem inicial. Dessa forma, soma-se os *partness map* para que sejam obtidas janelas da *face proposals* (YANG *et al.*, 2018).

Segundo Yang *et al.* (2018), a segunda etapa do Faceness-Net refina a janela de candidatos gerada a partir do primeiro estágio usando uma CNN multitarefa, onde a classificação de face e a *bounding box regression* são otimizadas em conjunto. Esta CNN multitarefa é pré-treinada nas camadas convolucionais de 1 a 4 do primeiro estágio. Na quarta camada, a CNN é dividida em 2 ramos, os quais consistem em 2 camadas convolucionais. Um destes ramos irá executar a classificação de face, e o outro fará a *bounding box regression* (YANG *et al.*, 2018).

Nos experimentos realizados, o método obteve bons resultados nos *datasets Annotated Faces in the Wild* (AFW) (ZHU; RAMANAN, 2012), *Pattern Analysis, Statistical Modelling and Computational Learning* (PASCAL) (NADA *et al.*, 2018), e WIDER FACE (YANG *et al.*, 2016). Sua acurácia no subconjunto de nível fácil foi de, respectivamente, 98,05%, 92,11% e 71,7% (YANG *et al.*, 2018).

De acordo com Yang *et al.* (2018), estas redes de reconhecimento de atributos podem compartilhar camadas de convolução iniciais, assim reduzindo em 83% os parâmetros, bem como as computações, além de melhorar a robustez para a detecção facial. Por fim, a eficácia do método também pode ser atingida a partir de atributos de imagens de faces não cortadas, sem supervisão explícita das partes.” (YANG *et al.*, 2018, p. 12). Além disso, eles também destacam que obtiveram sucesso ao introduzir o conceito de *faceness score*, o qual pode ser usado para ranquear janelas candidatas em qualquer técnica de proposta de região, podendo gerar conjuntos de propostas faciais com alta taxa de recall (YANG *et al.*, 2018).

2.2 DEFORMABLE FACE NET FOR POSE INVARIANT FACE RECOGNITION

Segundo He *et al.* (2019), a acurácia do reconhecimento facial tem aumentado rapidamente com o uso de *Convolutional Neural Network* (CNN). Entretanto, os autores ressaltam que sob circunstâncias não controladas, como poses, iluminação, resolução, oclusão e expressão, a performance do sistema de reconhecimento facial é

¹ Partness map: São os mapas de resposta que destacam 5 partes diferentes da face (YANG *et al.*, 2018, p. 1-2).

Excluído: de

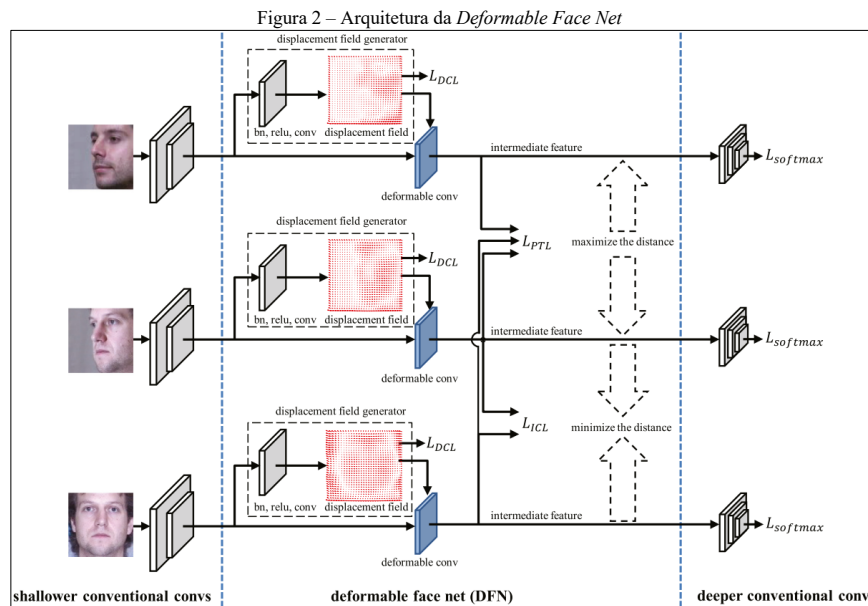
Excluído: , tradução nossa

Excluído: ¶

Excluído: , tradução nossa

afetada significativamente. Por conta disso, He *et al.* (2019) propuseram um método de alinhamento em nível de *feature* para lidar com variações de pose no reconhecimento de rosto.

He *et al.* (2019) ressaltam que o método inicialmente aplica uma *Deformable Face Net* (DFN), a qual busca aprender simultaneamente o alinhamento em nível de *feature* e a extração de *features* para reconhecimento de rosto por meio de convoluções deformáveis com um campo de deslocamento espacial. Dessa maneira, estes campos permitem que a DFN seja capaz de alinhar as *features*, independente de variações de pose. A Figura 2 apresenta a arquitetura da DFN.



Fonte: He *et al.* (2019).

A partir da Figura 2, pode-se observar que a DFN é composta de uma camada de convolução inicial com função de ativação *ReLU* e com a função de perda *Displacement Consistency Loss* (DCL) para que seja gerado um *displacement field*. Em seguida, passa por uma *deformable convolution*, a qual gera uma *feature* intermediária otimizada pelas funções de perda *Identity Consistency Loss* (ICL) e *Pose-Triplet Loss* (PTL). Por fim, as camadas mais profundas de convolução utilizam a função de perda *Softmax*.

Com o propósito de que os campos de deslocamento espacial sejam consistentes em sua orientação e amplitude, a DCL é utilizada para que obter um alinhamento melhor para reconhecer a face (HE *et al.*, 2019). Neste sentido, as funções ICL e PTL, servem para minimizar a variação de *features* intraclasse causada por poses diferentes e maximizar a distância de *recursos* entre classes sob poses semelhantes. Dessa forma, He *et al.* (2019) afirmam que a DFN supera os métodos de última geração nos seguintes *datasets* de testes: MegaFace (KEMELMACHER-SHLIZERMAN *et al.*, 2016), MultiPIE (GROSS *et al.*, 2008) e CFP (SENGUPTA *et al.*, 2016). No MegaFace, o método obteve 82,11% de precisão. No MultiPIE, obteve-se nas poses de $\pm 90^\circ$ uma acurácia de 85,66%. Por fim, o *dataset* CFP, obteve-se uma acurácia de 94,01%. Com isso, He *et al.* (2019) afirmam que o método aumenta significativamente o reconhecimento de rosto em todas as poses.

Excluído: 2019

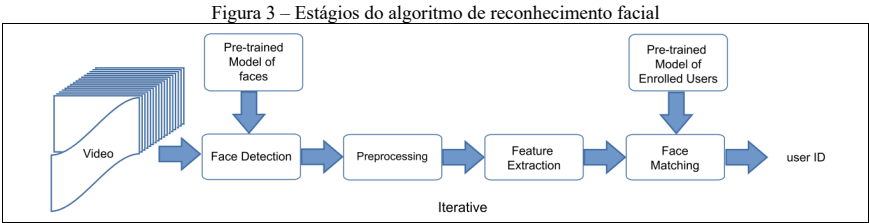
2.3 EFFICIENT FACE RECOGNITION SYSTEM FOR OPERATING IN UNCONSTRAINED ENVIRONMENTS

Sanchez-Moreno *et al.* (2021) desenvolveram um de sistema reconhecimento facial para ambientes irrestritos. Eles utilizaram o YOLO-Face modificado com um esquema de classificação que permite aumentar a acurácia da detecção com redução do tempo computacional. Os autores também propuseram uma arquitetura que combina uma CNN com um algoritmo de aprendizagem supervisionada, como por exemplo, o *Support Vector Machine* (SVM).

Segundo Sanchez-Moreno *et al.* (2021), subdividiu-se o algoritmo em 4 estágios. No primeiro estágio, realiza-se a detecção da face por meio de um *input* de *frame* de vídeo ou imagem. Para isso, eles utilizaram o

Excluído: :

YOLO-Face, ao qual detecta as faces em tempo real. Desse modo, os autores resolveram o problema de velocidade. A Figura 3 apresenta os estágios do algoritmo de reconhecimento facial.



Fonte: Sanchez-Moreno *et al.* (2021).

No segundo estágio, realiza-se o redimensionamento, normalização L2 e conversão da imagem para escala de cinza, tendo como intuito a preparação para aplicar o *machine learning model* (SANCHEZ-MORENO *et al.* 2021, p. 4-7). No terceiro estágio ocorre a extração das *features* através dos métodos de classificação como Support Vector Machine (SVM), Random Forest (RF) e K-Nearest Neighbors (KNN). Por fim, no último estágio, as *features* faciais são comparadas com as *features* conhecidas de usuários cadastrados para identificação pessoal.

Para fins de avaliação do reconhecimento facial, Sanchez-Moreno *et al.* (2021) utilizaram os seguintes *datasets*: *Labeled Faces in the Wild* (LFW) (HUANG *et al.*, 2007) e *YouTube Faces* (YTF) (WOLF; HASSNER; MAOZ, 2011). De acordo com os autores, no LFW, o método obteve uma acurácia de 99,1%. Já no YTF, a acurácia foi de 94,7%.

A partir desses resultados, Sanchez-Moreno *et al.* (2021) concluem que o sistema proposto melhora a precisão no banco de dados LFW, obtendo pontuação semelhante ou superior a diferentes estruturas de CNN utilizadas em reconhecimento faciais. Assim, no que diz respeito ao reconhecimento facial, os autores resumiram que “[...] com os avanços dos algoritmos de deep learning aliados aos avanços tecnológicos, é possível desenvolver sistemas de reconhecimento facial em tempo real capazes de operar em ambientes de alta complexidade.” (SANCHEZ-MORENO *et al.*, 2021).

3 PROPOSTA

Esta seção apresentada a justificativa para o desenvolvimento desse trabalho, os principais requisitos e a metodologia de desenvolvimento que será utilizada. Também são relacionados os assuntos e as fontes bibliográficas que irão fundamentar o estudo proposto.

3.1 JUSTIFICATIVA

No Quadro 1 é apresentado o comparativo entre os trabalhos correlatos. As linhas representam as características e as colunas os trabalhos.

Quadro 1 - Comparativo dos trabalhos correlatos			
Trabalhos correlatos	Yang <i>et al.</i> (2018)	He <i>et al.</i> (2019)	Sanchez-Moreno <i>et al.</i> (2021)
Características			
Detecção de faces	Sim	Sim	Sim
Reconhecimento de faces	Não	Não	Sim
Uso das regiões faciais para aprendizado	Sim	Não	Não
Lida com ambientes irrestritos (oclusões, iluminação, dimensionamento e poses)	Sim	Sim	Sim
Arquitetura de rede neural	Faceness-Net	DFN	YOLO-faces
Datasets utilizados	AFW, PASCAL, FDDB e WIDER FACE	MegaFace, MultiPIE e CFP	LFW e YTF
Taxa de acurácia para os datasets	98,05%, 92,11% e 71,7%	82,11%, 85,66% e 94,01%.	99,1% e 94,7%

Fonte: elaborado pelo autor.

A partir do Quadro 1, pode-se observar que todos os estudos utilizam de redes neurais convolucionais (CNN) como método de aprendizado de máquina. No que diz respeito a detecção de faces, todos os estudos correlatos implementam diferentes métodos. Yang *et al.* (2018) utilizaram atributos faciais para classificar e detectar faces. He *et al.* (2019) optaram pela rede *Deformable Face Net* (DFN), a qual simultaneamente alinha e

extraí as *features*. Já Sanchez-Moreno *et al.* (2021) utilizaram o método YOLO-faces, que calcula a probabilidade em cada região, destacando-as através de *bounding boxes*.

No que diz respeito ao reconhecimento facial, nota-se que apenas Sanchez-Moreno *et al.* (2021), implementam tal funcionalidade. Já Yang *et al.* (2018), com seu método Faceness-Net, utilizam regiões faciais, ou seja, cabelo, olhos, nariz, boca e barba para realizar a detecção facial. Já em relação aos ambientes irrestritos, isto é, levar em conta as oclusões, poses variadas, dimensões da imagem e iluminação, todos os estudos implementaram solução as quais conseguem obter bons resultados nestas condições adversas. Além disso pode-se observar que a acurácia obtida pelo método de Yang *et al.* (2018) nos *datasets* AFW, PASCAL, FDDB e WIDER FACE foi de 98,05%, 92,11% e 71,7%. Já as acurácias atingidas pelo método de He *et al.* (2019) nos *datasets* MegaFace, MultiPIE e CFP foram de 82,11%, 85,66% e 94,01%. E, Sanchez-Moreno *et al.* (2021) obtiveram 99,1% e 94,7% nos *datasets* LFW e YTF.

A partir deste contexto, este trabalho implementará um protótipo de detecção e reconhecimento facial em ambientes irrestritos para que, dado um conjunto de imagens, seja feita a pré-seleção automática para cada indivíduo informado. O protótipo a ser desenvolvido poderá ser utilizado por fotógrafos, aos quais após a realização de eventos como formaturas, casamentos e festas em geral, precisam realizar o processo de verificação de todas as imagens retiradas para encontrar e agrupar as fotos de acordo com quem aparece nas imagens. Ressalta-se que tal processo demanda muito tempo para ser feito manualmente, pois em um evento são tiradas milhares de fotos. Além disso, também se destaca a necessidade de estudos para definir qual é a qualidade mínima para o processamento e reconhecimento ocorrer com melhor precisão.

No cenário de fotografias profissionais, há alguns anos, o processo de envio das fotos tiradas pelo fotógrafo para que o cliente fizesse a seleção das mesmas era feito através de CDs ou DVDs, sendo posteriormente substituídos por *pendrives*. Deste modo, todos os fotógrafos que querem oferecer um serviço profissional e de qualidade para seu cliente, buscam formas de melhorar todos os processos que envolvem seu trabalho como um todo, e a tecnologia, técnicas de processamentos e aprendizado e máquina podem ajudá-los. Assim, este trabalho tem como essência a otimização do processo de seleção das fotografias, ao qual seus clientes apareçam nas imagens. Para isso, serão investigados e utilizados algoritmos de reconhecimento facial e redes neurais convolucionais que se adequem melhor a ambientes irrestritos.

3.2 REQUISITOS PRINCIPAIS DO PROBLEMA A SER TRABALHADO

Esta seção mostrará os Requisitos Funcionais (RF) e Requisitos Não Funcionais (RNF) para a implementação do protótipo de seleção de fotos a partir de faces. Os requisitos do trabalho estão divididos em duas partes: requisitos da aplicação web e requisitos do módulo de reconhecimento.

A aplicação web deverá:

- a) manter o cadastro de usuários (RF);
- b) permitir o login de usuários já cadastrados (RF);
- c) permitir ao usuário a criação de uma nova galeria de fotos (RF);
- d) permitir ao usuário visualizar as fotos pré-selecionadas (RF);
- e) utilizar o ambiente de desenvolvimento Visual Studio Code (RNF);
- f) desenvolver a aplicação web utilizando NodeJS no *backend* e React no *frontend* (RNF);
- g) utilizar o banco de dados PostgreSQL para persistir os dados (RNF).

O módulo de reconhecimento deverá:

- a) utilizar técnicas de processamento de imagens para melhorar a qualidade das imagens obtidas (filtragem de ruído e normalização da iluminação) (RF);
- b) realizar a normalização e a extração de características para estabelecer a qualidade mínima necessária da imagem para realizar a detecção e o reconhecimento facial (RF);
- c) utilizar algoritmos de aprendizado de máquina para efetuar a detecção das faces e consecutivamente a seleção das fotos a partir da face dos indivíduos (RF);
- d) utilizar métricas de similaridade para comparar as faces de entrada com os presentes nas fotos (RF);
- e) utilizar métricas de classificação: PCA, LPA, ICA, SVM, RF, KNN ou identificar outro classificador que produza melhor resultado (RF);
- f) ser desenvolvida na linguagem de programação Python (RNF);
- g) utilizar as bibliotecas OpenCV, Keras e scikit-learn (RNF).

3.3 METODOLOGIA

O trabalho será desenvolvido observando as seguintes etapas:

- a) levantamento bibliográfico: pesquisar sobre técnicas e novas tecnologias para reconhecimento facial de pessoal, processamento de imagens para melhoria no reconhecimento, aprendizado de máquina e

Formatado: Fonte: Itálico

Formatado: Fonte: Itálico

- trabalhos correlatos;
- elicitação dos requisitos da aplicação web: baseando-se nas informações da etapa anterior, reavaliar os requisitos propostos para a aplicação, tanto funcionais como não funcionais;
 - especificação: utilizar a ferramenta de diagramação Enterprise Architect (EA) para elaborar os diagramas de caso de uso e de atividades de acordo com a Unified Modeling Language (UML);
 - implementação: a partir do item (c) implementar a aplicação web utilizando o NodeJS no *backend* e React no *frontend*;
 - testes da aplicação web: elaborar testes para validar a usabilidade da aplicação web;
 - coleta e montagem da base de dados: coletar fotos/vídeos ou utilizar bases de dados com faces obtidas/cedidas junto a projetos já existentes na comunidade acadêmica;
 - preparação das imagens: realizar o realce e melhoramento das imagens utilizando o OpenCV. Posteriormente subdividir as imagens em dois grupos: imagens de treinamento e de validação;
 - definição da arquitetura da rede neural convolucional: definir a arquitetura da rede neural convolucional mais aderente ao processo de reconhecimento facial;
 - implementação do reconhecimento facial: a partir do item (h) realizar a implementação do módulo de reconhecimento facial utilizando OpenCV para o processamento das imagens e a biblioteca Keras para aprendizado de máquina;
 - testes do reconhecimento facial: paralelamente à implementação do item (i), realizar os testes do módulo de reconhecimento facial, com base em um banco de imagens de familiares e amigos.

As etapas serão realizadas nos períodos relacionados no Quadro 2.

Quadro 2 – Cronograma de atividades a serem realizadas

etapas / quinzenas	2023											
	jan.		fev.		mar.		abr.		maio		jun.	
	1	2	1	2	1	2	1	2	1	2	1	2
levantamento bibliográfico												
elicitação de requisitos da aplicação web												
especificação												
implementação												
testes da aplicação web												
coleta e montagem da base de dados												
preparação das imagens												
definição da arquitetura da rede neural convolucional												
implementação do reconhecimento facial												
testes do reconhecimento facial												

Fonte: elaborado pelo autor.

4 REVISÃO BIBLIOGRÁFICA

Nesta seção encontra-se uma breve descrição dos assuntos relacionados a este estudo: Reconhecimento facial, visão computacional e Rede Neural Convolucional.

O reconhecimento facial é uma técnica de biometria baseada nos traços do rosto das pessoas. Não é segredo que os seres humanos têm uma facilidade para reconhecer os rostos das pessoas com precisão. Mesmo que duas pessoas sejam parecidas, nós somos capazes de distingui-las. Essa técnica pode ser transmitida para os computadores, mas, para isso, precisa-se definir traços únicos que devem ser mapeados em códigos binários. (PISA, 2022).

Segundo Sanchez-Moreno *et al.* (2021), os sistemas de reconhecimento facial, devem primeiro detectar uma face independente do meio em que se encontra, para que finalmente possa comparar as *features* extraídas com as faces cadastradas previamente, assim fazendo a correspondência de face. Ainda de acordo com os autores, uma das aplicações para o reconhecimento facial é seu uso em sistemas de segurança em tempo real. Além disso, Sanchez-Moreno *et al.* (2021) também afirmam que os algoritmos ou métodos que utilizam inteligência artificial, mais especificamente redes neurais, emulam a capacidade do ser humano de reconhecer uma imensa variedade de objetos, animais, faces, etc.

Segundo Rosa (2018), a Rede Neural Convolucional (do inglês *Convolutional Neural Network* - CNN) é uma classe da rede neural artificial que começou a ser aplicada no processamento e análise de imagens digitais. A CNN é uma variante da rede *perceptrons* de múltiplas camadas, e foi inspirada no processo biológico de processamentos de dados visuais. As células responsáveis pelo processamento dos dados visuais são sensíveis a pequenas sub-regiões do campo visual, também chamado de campo receptivo. Essas células são adequadas para explorar a correlação espacial local presente em imagens naturais, visto que elas atuam como filtros locais sobre o campo de visão. O campo de visão é todo coberto através dos campos receptivos de diferentes neurônios que se sobrepõem parcialmente. Imagens coloridas em treinamentos podem resultar numa melhora significativa na

Excluído: .

Excluído: .

interpretação dos dados, pois a CNN tira maior proveito nestas imagens do que em imagens em escala de cinza. Isto ocorre porque assim como nos humanos, a cor ajuda na análise e interpretação de objetos, facilitando a memorização e reconhecimento.

REFERÊNCIAS

GROSS, Ralph *et al.*, Multi-PIE. In: IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE & GESTURE RECOGNITION, 8., 2008, Amsterdam. **Proceedings...** Amsterdam, 2008. p. 1-8.

HE, Mingjie *et al.*, Deformable Face Net: Learning Pose Invariant Feature with Pose Aware Feature Alignment for Face Recognition. In: IEEE INTERNATIONAL CONFERENCE ON AUTOMATIC FACE & GESTURE RECOGNITION, 14., 2019, Lille. **Proceedings...** Lille: Lille grand Palais, 2019. p. 1-8.

HUANG, Gary B. *et al.* **Labeled Faces in the Wild**: A Database for Studying Face Recognition in Unconstrained Environments. 2007. 11 f. Technical Report 07-49 - University of Massachusetts, Amherst.

KEMELMACHER-SHLIZERMAN, Ira *et al.*, **The MegaFace Benchmark**: 1 Million Faces for Recognition at Scale. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 29., 2016, Las Vegas. **Proceedings...** Las Vegas: Caesars Palace, 2016. p. 4873-4882.

NADA, Hajime *et al.*, **Pushing the Limits of Unconstrained Face Detection**: a Challenge Dataset and Baseline Results. 2018. Disponível em: <https://arxiv.org/abs/1804.10275>. Acesso em: 27 set. 2022

PISA, Pedro. **Como funciona o reconhecimento facial**. Disponível em: <http://www.techtudo.com.br/artigos/noticia/2012/04/como-funciona-o-reconhecimento-facial.html>. Acessado em: 13 set. 2022

ROSA, Renan P. **Método de classificação de pragas por meio de rede neural convolucional profunda**. 2018. 101 f. Dissertação (Mestrado em Computação Aplicada) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2018.

SANCHEZ-MORENO, Alejandra S. *et al.* 2021. Efficient Face Recognition System for Operating in Unconstrained Environments. **Journal of Imaging**, v. 7, n. 9, p. 161, ago. 2021.

SENGUPTA, Soumyadip *et al.*, Frontal to profile face verification in the wild. In: IEEE WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION, 16., 2016, Lake Placid. **Proceedings...** Lake Placid: Crowne Plaza, 2016. p. 1-9.

WOLF, Lior; HASSNER, Tal; MAOZ, Itay. Face Recognition in Unconstrained Videos with Matched Background Similarity. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 24., 2011, Colorado Springs. **Proceedings...** Colorado Springs: Crowne Plaza, 2011. p. 529-534.

YANG, Shuo *et al.* Faceness-Net: Face Detection through Deep Facial Part Responses. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 40, n. 8, p. 1845-1859, ago. 2018.

YANG, Shuo *et al.*, **Wider Face**: A Face Detection Benchmark. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 29., 2016, Las Vegas. **Proceedings...** Las Vegas: Caesars Palace, 2016. p. 5525-5533.

ZHU, Xiangxin; RAMANAN, Deva. Face detection, Pose Estimation, and Landmark Localization in the Wild. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 25., 2012, Providence. **Proceedings...** Providence: Rhode Island Convention Center, 2012. p. 2879-2886.

Excluído: <

Excluído: >

Excluído: <

Excluído: >

FORMULÁRIO DE AVALIAÇÃO BCC – PROFESSOR TCC I – PRÉ-PROJETO

Avaliador(a): Dalton Solano dos Reis

ASPECTOS AVALIADOS		atende	atende parcialmente	não atende
ASPECTOS TÉCNICOS	1. INTRODUÇÃO O tema de pesquisa está devidamente contextualizado/delimitado?	X		
	O problema está claramente formulado?	X		
	2. OBJETIVOS O objetivo principal está claramente definido e é passível de ser alcançado?	X		
	Os objetivos específicos são coerentes com o objetivo principal?	X		
	3. JUSTIFICATIVA São apresentados argumentos científicos, técnicos ou metodológicos que justificam a proposta?	X		
	São apresentadas as contribuições teóricas, práticas ou sociais que justificam a proposta?	X		
	4. METODOLOGIA Foram relacionadas todas as etapas necessárias para o desenvolvimento do TCC?	X		
	Os métodos, recursos e o cronograma estão devidamente apresentados?	X		
ASPECTOS METODOLÓGICOS	5. REVISÃO BIBLIOGRÁFICA Os assuntos apresentados são suficientes e têm relação com o tema do TCC?	X		
	6. LINGUAGEM USADA (redação) O texto completo é coerente e redigido corretamente em língua portuguesa, usando linguagem formal/científica?	X		
	A exposição do assunto é ordenada (as ideias estão bem encadeadas e a linguagem utilizada é clara)?	X		
	7. ORGANIZAÇÃO E APRESENTAÇÃO GRÁFICA DO TEXTO A organização e apresentação dos capítulos, seções, subseções e parágrafos estão de acordo com o modelo estabelecido?	X		
	8. ILUSTRAÇÕES (figuras, quadros, tabelas) As ilustrações são legíveis e obedecem às normas da ABNT?	X		
	9. REFERÊNCIAS E CITAÇÕES As referências obedecem às normas da ABNT?	X		
	As citações obedecem às normas da ABNT?	X		
Todos os documentos citados foram referenciados e vice-versa, isto é, as citações e referências são consistentes?		X		



ATA DA DEFESA: BANCA DO PRÉ-PROJETO

Venho, por meio deste, manifestar minha avaliação sobre a **apresentação** do Pré-Projeto de TCC realizado pelo(a) acadêmico(a), **_ Jadiel dos Santos _** no **SEGUNDO SEMESTRE DE 2022**, com o título **_ PROTÓTIPO DE VISÃO COMPUTACIONAL PARA SELEÇÃO DE FOTOS COM FACES _**, sob orientação do prof(a). **_ Aurélio Faustino Hoppe _**.

A referida apresentação obteve a seguinte nota:

Componente da Banca	Nota (de 0 a 10)
Professor(a) Avaliador(a): _ Andreza Sartori _	10

ATENÇÃO. A nota acima se refere somente a apresentação do pré-projeto e vai ser repassada para o aluno (orientando). Favor preencher os campos acima e enviar por e-mail ao professor de TCC1 (dalton@furb.br). Não passar o arquivo com as anotações da revisão já enviado ao professor de TCC1 para o orientando e nem para o professor orientador. Após o professor de TCC1 receber esta ata preenchida, o professor de TCC1 vai disponibilizar para o orientando/orientador os arquivos com as revisões. Caso julgue necessário fazer mais alguma consideração relacionada ao pré-projeto ou a defesa, favor usar o espaço abaixo.

Observações da apresentação: O trabalho foi bem apresentado, porém o aluno deverá fazer as devidas correções no trabalho escrito.



ATA DA DEFESA: BANCA DO PRÉ-PROJETO

Venho, por meio deste, manifestar minha avaliação sobre a **apresentação** do Pré-Projeto de TCC realizado pelo(a) acadêmico(a), **_ Jadiel dos Santos _** no **SEGUNDO SEMESTRE DE 2022**, com o título **_ PROTÓTIPO DE VISÃO COMPUTACIONAL PARA SELEÇÃO DE FOTOS COM FACES _**.

A referida apresentação obteve a seguinte nota:

Componente da Banca	Nota (de 0 a 10)
Professor(a) Orientador(a): _ Aurélio Faustino Hoppe _	10

A apresentação aconteceu em **_07_ / _11_ / 2022** na sala de reunião virtual do MS-Teams, tendo início às **_17_:30_** hs e foi encerrada às **_18_:00_** hs.

ATENÇÃO. A nota acima se refere somente a apresentação do pré-projeto e vai ser repassada para o aluno (orientando). Favor preencher os campos acima e enviar por e-mail ao professor de TCC1 (dalton@furb.br). Lembro que os arquivos com as anotações das revisões do professor de TCC1 e Avaliador serão enviados para o orientando e professor orientador após o professor de TCC1 receber esta ata preenchida. Caso julgue necessário fazer mais alguma consideração relacionada ao pré-projeto ou a defesa, favor usar o espaço abaixo.

Observações da apresentação: