SISTEMA DE RECONHECIMENTO E IDENTIFICAÇÃO FACIAL EM DISPOSITIVO EMBARCADO

DIOGO HONORATO MONTAGNER

SÃO PAULO

2020

SISTEMA DE RECONHECIMENTO E IDENTIFICAÇÃO FACIAL EM DISPOSITIVO EMBARCADO

DIOGO HONORATO MONTAGNER

Monografia apresentada como requisito à obtenção do título de Tecnólogo em Eletrônica Industrial, Curso de Tecnologia em Eletrônica Industrial, Departamento de Sistemas Eletrônicos, Faculdade de Tecnologia de São Paulo.

Orientador: Prof. Dr.Victor Sonnenberg

SÃO PAULO

2020

Montagner, Diogo Honorato

Sistema de reconhecimento e identificação facial em dispositivo embarcado / Diogo Honorato Montagner. São Paulo, 2020.

Orientador: Dr. Victor Sonnenberg.

Trabalho de Conclusão de Curso para obtenção do título de Tecnólogo em Eletrônica Industrial – Faculdade de Tecnologia de São Paulo.

Inclui referência.

1. Visão Computacional. 2. Raspberry Pi. 3. OpenCV. I. Sistema de reconhecimento e identificação facial em dispositivo embarcado. II. Montagner, Diogo Honorato. III. Faculdade de Tecnologia de São Paulo.

Agradecimentos

Agradeço aos meus professores pela dedicação e paciência nessa caminhada que é a graduação, em especial ao professor Sonnenberg, que aceitou o desafio de orientar meu trabalho. Agradeço também aos professores Leonardo Frois, Maurício Deffert, Roberto Katsuhiro, Ricardo Rangel, Aparecido Nicolett e todos os demais professores que compõe a equipe docente da FATEC São Paulo.

Não posso deixar de agradecer à todos os profissionais do Instituto de Pesquisas Tecnológicas do Estado de São Paulo (IPT), instituto em que trabalho, que me proporcionaram conhecimento e incentivo para meu aprimoramento pessoal e profissional, parte desse conhecimento expresso nesse trabalho de conclusão de curso. Agradeço em especial à Felipe e Celso pelas dicas e consultorias no aprendizado das técnicas de visão computacional. Agradeço também à Rodrigo e Matheus que aparecem em algumas figuras desse trabalho como “Superman” e “Clark do SmallVille” respectivamente.

Agradeço a todos os meus parentes que de alguma maneira contribuíram, ou não, para esse trabalho, a começar pelas minhas irmãs Nat e Carol, aos meus queridos Thiago, Elisa e Poketânia, a Cezar e Lívia, ao Preto Rei, a todos os meus parceiros de graduação.

Para ela, um parágrafo a parte, afinal, quem faz questão de estar lado a lado com você, tanto nos momentos de força, quando nos momentos de fraqueza, faz jus a um destaque especial que, dentro das minhas limitações, não conseguirei descrever, tamanha importância que cumpre na minha vida, Giovanna sobrenome.

LISTA DE FIGURAS

RESUMO

No terreno do que se convencionou chamar Inteligência Artificial (AI) foram desenvolvidas muitas tecnologias, um ramo dessas tecnologias é chamado *Computer Vision*, em português Visão Computacional. Uma vez definido esse ramo da AI, junto da *Computer Vision* surgiram também possibilidades até então inalcançadas por máquinas, como a capacidade de "ver", assim como os humanos veem, porém, com uma capacidade muito superior de processamentos. Não obstante, vivemos ainda a era da Indústria 4.0, conceito definido em função das infraestruturas tecnológicas representadas por inúmeros sensores e computadores conectados entre si pela grande rede mundial, a internet, gerando ainda um outro conceito reproduzido amplamente por muitos palestrantes e divulgadores de tecnologias, o conceito de *Internet of Things* (IoT), a Internet das Coisas. Todos esses dispositivos e sensores alimentam grandes bancos de dados, daí advém o que chamamos de *Big Data*, de onde é possível extrair conhecimento e padrões que, vistos a partir de um elemento individual, não se consegue notar mas, a partir dessa grande quantidade de informação, pode-se tirar grande proveito. O presente trabalho propõe a construção de um sistema de reconhecimento e identificação facial usando um Raspberry Pi para processamento das imagens como proposta barata e eficiente para algumas soluções que serão arroladas. O sistema está funcional, porém ocorrem falsos positivos e reconhecimentos de objetos como faces. Sugestões de soluções aos problemas encontrados também foram abordados e propostas apontadas.

Palavras-chave: Raspberry Pi, Visão Computacional, Inteligência Artificial, Reconhecimento Facial, OpenCV.

ABSTRACT

In the field of what was conventionally called Artificial Intelligence (AI), many technologies were developed, a branch of these technologies is called Computer Vision, in Portuguese Computer Vision. Once this branch of AI was defined, with Computer Vision there were also possibilities hitherto unreached by machines, such as the ability to “see” as humans see, however, with a much higher processing capacity. Nevertheless, we still live in the era of Industry 4.0, a concept defined by the technological infrastructures represented by countless sensors and computers connected to each other by the great world network, the internet, generating yet another concept widely reproduced by many speakers and disseminators of technologies, the concept of Internet of Things (IoT), the Internet of Things. All of these devices and sensors feed large databases, hence what we call Big Data, from which it is possible to extract knowledge and standards that, seen from an individual element, cannot be noticed, but from that large amount of information, great advantage can be taken. The present work proposes the construction of a facial recognition and identification system using a Raspberry Pi for image processing as a cheap and efficient proposal for some solutions that will be listed. The system is functional, but false positives and recognitions of objects such as faces occur. Suggestions for solutions to the problems encountered were also addressed and solutions pointed out.

Key-words: Raspberry Pi, Computer Vision, Artificial Intelligence, Face Recognition, OpenCV.

SUMÁRIO

1. INTRODUÇÃO

Na crescente produção de dispositivos computacionais cada vez mais baratos e com maior poder de processamento e, frente às tecnologias que flertam com o que é chamado de Inteligência Artificial (AI), o presente trabalho tem a intenção de apresentar um projeto simples e barato de um sistema de detecção e identificação facial.

O uso de câmeras eletrônicas para os mais diversos fins está presente no dia a dia de todos. A primeira câmera de segurança foi instalada na Alemanha, para ver o lançamento do foguete V2. Os primeiros registros do uso de câmeras em ambiente público foram feitos pela polícia de New York que instalou câmeras para a vigilância policial das ruas(REVISTA SEGURANÇA ELETRÔNICA). Com o tempo, as câmeras eletrônicas começaram a ganhar os setores privados, principalmente para uso em ambientes que é exigido segurança como bancos e penitenciarias. Não tardou para elas ganharem ampla popularidade nos setores privado e público a ponto de se tornarem indispensáveis, inclusive a existência de câmeras tornou-se requisito para qualidade de segurança no monitoramento de ambientes externos e internos de condomínios residenciais, industriais e empresariais.

Com o tempo e, combinado com o avanço tecnológico, as câmeras que até então transmitiam os dados captados por cabo, hoje transmitem as imagens via rede wireless, ou mesmo conectadas à internet, com as chamadas *Cameras Internet Protocol* (Câmeras IPs), câmeras que são conectados a uma rede e podem ter seus dados de imagens acessados por qualquer dispositivo conectado na mesma rede(REVISTA DIGITAL SECURITY).

Com o advento dos smartphones e das redes sociais digitais, as câmeras estabeleceram na vida das pessoas novos hábitos e dependências. Vemos com frequência imagens captadas por essas câmeras nos noticiários mostrando a ocorrência de alguns crimes diversos, hoje todos temos uma “câmera de vigilância” na mão que pode registrar imagens dos crimes, por exemplo, um ladrão que invade um estabelecimento, um sequestro ou alguém furtando um produto de determinada loja. Não há como escapar das câmeras, elas tomam conta da vida das pessoas.

Que as câmeras fazem parte do nosso cotidiano isso é inegável, o questionamento que deve ser levantado é se as imagens que estão sendo tomadas nos diversos ambientes

têm realmente um bom aproveitamento. Um sistema de circuito fechado de televisão (CFTV) precisa de pessoas para monitorar as imagens projetadas nas telas, porém, é sabido que pessoas que trabalham muito tempo ao longo de jornadas extensas tem sua atenção diminuída. Também é comum os estabelecimentos comerciais terem câmeras externas e internas, mas sem ninguém monitorando as imagens e estas apenas ficam gravadas num banco de dados para consulta caso haja alguma ocorrência. Se essa ocorrência fosse, por exemplo, um assalto decorrente de um suspeito que ficava à espreita, ao redor desse estabelecimento, seria possível que esse crime fosse prevenido caso houvesse uma melhor apreciação das imagens externas desse local.

Até aqui levantamos alguns usos das câmeras no setor da segurança, mas há muitas outras aplicações que podem ser pensadas para o uso delas, como por exemplo, o monitoramento de desmatamento de regiões florestais. O Instituto Nacional de Pesquisas Espaciais (INPE) é especialista em monitorar desmatamento e queimadas das florestas brasileiras através de tratamento de imagens de câmeras de satélites. Existem ainda câmeras especiais que captam ondas de infravermelho, característica importante quando é preciso monitorar a temperatura de alguém, pois essas câmeras podem nos dar imagens de pessoas com febre que, num contexto de pandemia como do COVID-19, saber quem está ou não com febre apenas visualizando uma imagem pode ser mais eficiente que muitos sensores de aferição de temperatura(OLIVEIRA, 2020). Além disso, as câmeras de infravermelho também podem ser usadas para monitorar equipamentos elétricos, a fim de identificar grande consumo de potência e prevenindo que circuitos sejam queimados por excesso de gasto de corrente elétrica.

Todas as informações referentes à segurança, ou temperatura, ou qualquer outra característica que seja importante extrair de uma imagem são feitas por pessoas, um ser humano que olha a imagem e vê nessa imagem algo que lhe é pertinente. Para ver se alguém está com febre numa imagem de câmera de infravermelha, como comentado acima, é preciso que alguém analise essa imagem mas, como já dito, funcionários após uma jornada extenuante de trabalho podem apresentar falta de concentração e deixar passar alguém com febre, possivelmente contaminado, entrando em contato com outras pessoas saudáveis devido a essa falta de atenção. Podemos prevenir essas falhas que podem acometer qualquer trabalhador com o auxílio de ferramentas computacionais. É

possível o próprio sistema identificar e interpretar as imagens das câmeras infravermelho emitindo um alerta quando alguém febril fosse notado na captura da câmera pois, felizmente a computação tem feito avanços significativos no tratamento e interpretação das imagens de câmeras, atribuindo a elas uma inteligência artificial capaz de “ver” melhor que os olhos humanos.

Essa característica da máquina “enxergar” é um campo da inteligência artificial conhecido como Visão Computacional. Essa habilidade de “ver”, de usar as câmeras como olhos para as máquinas, são de grande valia para pensar em soluções e otimizações de trabalhos que envolvam captura e interpretações de imagens.

É muito comum encontrar sistemas de vigilância de CFTV que precisam de cabeamento para conectar as câmeras numa central, numa máquina para armazenar as imagens capturadas e, dependendo do contexto, é necessária uma sala com telas e profissionais para monitorar os ambientes em tempo real. O uso de Inteligência Artificial pode ser útil para melhorar de alguma forma esse contexto.

1.1 **OBJETIVOS**

1.1.1 OBJETIVO GERAL

Criar um sistema de monitoramento dotado de inteligência artificial para detecção e identificação facial em hardware de baixo custo e baixo poder de processamento.

1.1.2 OBJETIVO ESPECÍFICO

1) Aquisição do hardware de baixo custo e baixo poder de processamento;
2) Construção de base de dados para identificação facial;
3) Treinamento de algoritmo de inteligência artificial;
4) Construção do sistema de detecção e identificação facial.

2. REVISÃO TEÓRICA

Entender as ideias e tecnologias por trás de um sistema de detecção e identificação facial é fundamental para entender suas potencialidades e limitações.

2.1 Raspberry Pi

O Rapsberry Pi (RPi) é considerado o computador mais barato do mundo (MERCES, 2014). Esse computador, que vem ganhando grande popularidade, foi lançado no fim de fevereiro de 2012 e a proposta inicial dos membros do Laboratório de Computação da Universidade de Cambridge, no Reino Unido, era de criar um computador barato e acessível para estimular jovens ao aprendizado da programação e, com isso, produzir projetos criativos para abrir portas a possíveis soluções de mercado.

Suas dimensões físicas impressionam logo de cara, pois se aproxima ao tamanho de um cartão de crédito, um pouco mais volumoso. Atualmente o RPi se encontra na sua versão Raspberry Pi 4. Além de ser um computador completo e barato, o Raspberry Pi possui algumas peculiaridades interessantes. Na sua plataforma existem 40 pinos de Input/Output, permitindo que sejam conectados periféricos como sensores e outros dispositivos, assim o RPi atua como um microcontrolador, podendo ser programado em linguagem de programação Python.

Outra peculiaridade interessante que vem sendo explorada é a sua funcionalidade como servidor. Bastam poucas e simples configurações para que o RPi se transforme num servidor podendo suportar muitas aplicações e serviços. Uma vez configurado como servidor, basta o cabo de alimentação para energizar o RPi e um cabo no conector *Gigabit Ethernet* para o servidor estar conectado à rede, porém o cabo também pode ser dispensado, uma vez que o RPi pode se conectar a rede via wireless. Esse formato dispensa uso de monitor, teclado, mouse e outros periféricos possíveis de nele estarem conectados. É possível acessar o console do RPi por outra máquina, via interface de comunicação disponível no portal raspberrypi.org, assim, mesmo sem monitor, sem teclado, o usuário consegue acessar e manipular configurações e diretórios do RPi.

2.2 IoT (“Internet of Things”)

Vamos imaginar um agricultor que recebe em seu smartphone informações em tempo real sobre as condições da sua plantação, se está sendo atacada por pragas, se a terra está precisando ser adubada, com todas essas informações combinadas com as previsões meteorológicas, o agricultor tem, por consequência, as melhores sugestões de colheitas e manutenção que o algoritmo de inteligência artificial pode ofertar a partir desses dados.

Se um corredor sob supervisão de um *personal trainer* é dotado de dispositivos no tênis, na palmilha e em outras partes do corpo, o *personal trainer* consegue monitorar os indicadores do corredor como atividade cardíaca, distância percorrida, calorias gastas em tempo real.

Na indústria, máquinas e sensores se comunicam de forma instantânea informando a necessidade de abastecimento na linha de produção, ou mesmo informam a necessidade de regularização de Ph num determinado material, e ainda, o atraso para terminar o processo de produção decorrente de algum imprevisto na linha.

As situações citadas acima são exemplos de aplicações de *Internet of Things* (IoT), em português a Internet das Coisas. Este conceito ainda não é fechado pois, muitos autores ainda divergem em alguns pontos, mas já é possível esboçar as ideias que circundam esse termo tão utilizado nas empresas e tão presente na boca de intelectuais que lidam com tecnologia. A comunicação de humano para humano, humano para máquina e mesmo da máquina para máquina, que hoje acontece por sensores e dispositivos, todos conectados a internet, trocando informações em tempo real, alimentando grandes bancos de dados têm caminhado para um ecossistema completamente diferente de comunicação e tomadas de decisão como jamais havíamos visto. De certo, muitas vantagens serão extraídas desse novo cenário e novos hábitos para as pessoas e indústrias serão esculpidos (MAGRANI, 2018).

2.3 Inteligência Artificial

Ainda com definição abrangente, alguns autores afirmam que AI é o estudo das faculdades mentais aplicados a modelos computacionais, outros dizem que AI é o

estudo de como os computadores podem realizar tarefas humanas e, ainda, há quem diga que AI é o estudo de como o computador irá perceber, raciocinar e agir. Sem contar aqueles que definem AI como sendo máquinas que através de sensores captam o que acontece no meio em que está inserido e, por meio de atuadores, agem sobre esse meio (SILVA NETO, 2016).

Podemos afirmar que AI é o campo da computação que busca reproduzir a capacidade de expressão cognitiva do ser humano, reproduzindo aspectos como raciocínio, percepção, aprendizagem e adaptação, se fazendo valer de matemática de probabilidade, modelagem heurística e ferramentas que flertam com esse campo complexo da computação.

Mesmo que a ideia de Inteligência Artificial seja antiga, o grande precursor dessa temática foi o Alan Mathison Turing que, ao publicar o artigo *Computing Machinery and Intteligence* divulgado na revista *Mind*, onde descreve o "Teste de Turing", inaugura todo um terreno fértil ao cultivo da Inteligência Artificial.

2.4 Visão Computacional

Ballard e Brown já discutiam em sua obra *Computer Vision*, no ano de 1982, o conceito de Visão Computacional, definindo esta como sendo a ciência que estuda e desenvolve tecnologia para aquisição de imagens por diferentes sensores e dispositivos a fim de tratar e interpretar tais imagens (BARELLI, 2018).

Em analogia a visão biológica, os nossos olhos são os dispositivos capazes de captar as imagens ao nosso redor, as imagens captadas são transformadas em sinais elétricos e enviadas ao nosso cérebro para que sejam tratadas e interpretadas. Na Visão Computacional, câmeras são os dispositivos capazes de captar as imagens ao redor, transformando-as em linguagem de máquina e, após essa transformação, é feito um tratamento desse sinal pela máquina. Os avanços e desenvolvimentos das bibliotecas para esse fim tem contribuído muito para que a máquina entenda e interprete o meio externo através das imagens. As máquinas "enxergam", assim como os humanos, porém, com capacidade de interpretação e processamento muito superior.

3. DESENVOLVIMENTO

A construção do sistema para atender as expectativas do trabalho será em cima de uma plataforma do tipo Raspberry Pi 3 Model B+, um computador simples e de baixo custo, porém adequado ao objetivo proposto. É possível ver a imagem desse modelo de computador na Figura 1.

Figura 1 – Imagem de um hardware Raspberry Pi 3 Model B+

Fonte: Página do Raspberry Pi[1]

Um aspecto interessante a ser destacado nesse sistema é que o processamento das imagens capturadas para a detecção e identificação facial acontece no próprio Raspberry Pi, dando a esse sistema a característica de Computação de Borda, que significa que os dados têm o processamento da imagem acontecendo no próprio Raspberry Pi.

O próximo passo é entender em que ambiente será desenvolvido a programação. No sistema operacional do Raspberry Pi, o *Raspbian*, há alguns *Integrated Development Environment* (IDE) nativos, ambientes de programação já embutidos no sistema operacional, um deles é o *Integrated Development and Learning Environment* (IDLE).

Esse é um ambiente de desenvolvimento em linguagem Python e será nesse ambiente, com essa linguagem, que será feito o desenvolvimento da programação.

Para a estruturação do programa foi usado uma biblioteca *open source* popularmente usada para aplicações de visão computacional chamada OpenCV. Essa biblioteca dispõe de muitas ferramentas e algoritmos para o emprego de técnicas de detecção e identificação facial. Dos recursos disponíveis, dois deles são fundamentais para o sucesso do sistema, são eles o algoritmo HaarCascade e o *Local Binary Patterns Histograms* (LBPH), um para a detecção facial e outro para a identificação facial, respectivamente.

3.1 Raspbian

Raspbian é um sistema operacional livre e gratuito. É derivado do Debian, sistema operacional baseado em kernel Linux, também gratuito. O Raspbian é uma adaptação para o hardware Raspberry Pi, e ele vem com mais de 35000 pacotes pré-compilados em formato agradável e fácil de serem instalados (DEBIAN PROJECT).

3.2 IDLE

Integrated Development and Learning Environment (IDLE) é o editor de texto para codificação em Python puro. Tem característica multiplataforma, funcionando praticamente da mesma forma tanto em Unix, Windows e macOS. Seu interpretador interativo, a janela de shell, marca em colorido a entrada de código, a saída e mensagens de erro. É possível abrir múltiplas janelas, tem recuo inteligente, dicas de chamada, preenchimento automático de código, e outras ferramentas que auxiliam no desenvolvimento da programação (PYTHON SOFTWARE FUNDATION).

3.3 Python

Python é uma linguagem de programação orientada a objetos fácil de usar e vem ganhando muita popularidade entre os desenvolvedores. Possui grande quantidade de bibliotecas, dando suporte a tarefas de programações comuns como, conexões de servidores web, buscas com expressões regulares e outras.

Usando uma sintaxe elegante, Python torna os programas escritos mais fáceis de serem compreendidos.

Os tipos de dados são fortes e dinamicamente tipados, que significa que suas variáveis não precisam ser previamente definidas, elas recebem suas definições durante a execução da aplicação, diferente de outras linguagem como Java ou C# onde a variável necessita de definição, precisa ser tipada, antes de ser usada. Python pode ser modificado e distribuído livremente pois, embora a linguagem tenha direitos autorais, ela está inscrita sob uma licença de código aberto (PYTHON.ORG).

3.4 OpenCV

Open Source Computer Vision Library (OpenCV) é uma biblioteca para software e sistemas de visão computacional e aprendizado de máquina. Por ser de licença *Berkeley Software Distribution* (BSD), é uma biblioteca de código aberto. Possui mais de 2500 algoritmos para aplicações de visão computacional e aprendizado de máquina. Amplamente usado por diversas empresas como Google, Yahoo, Microsoft, Intel, Sony, Honda, Toyota e por diversas *startups*. Os algoritmos dessa biblioteca vem sendo empregados para a detecção e identificação facial, identificação de objetos, rastreamento de movimento de câmera, rastreamento de objetos em movimento, extração 3D de objetos, junção de imagens para produção de imagens de alta resolução, remoção de olhos vermelhos de imagens tiradas com flash, reconhecimento de cenário e estabelecimento de marcadores para sobreposição com realidade aumentada. Orbitam aproximadamente 47 mil usuários em torno do OpenCV e é estimado um número de downloads na ordem de 18 milhões (OPENCV TEAM).

Essa biblioteca possui interface para C++, Python, Java, MATLAB, todas essas suportadas em Windows, Linux, Android e macOS.

3.5 HaarCascade

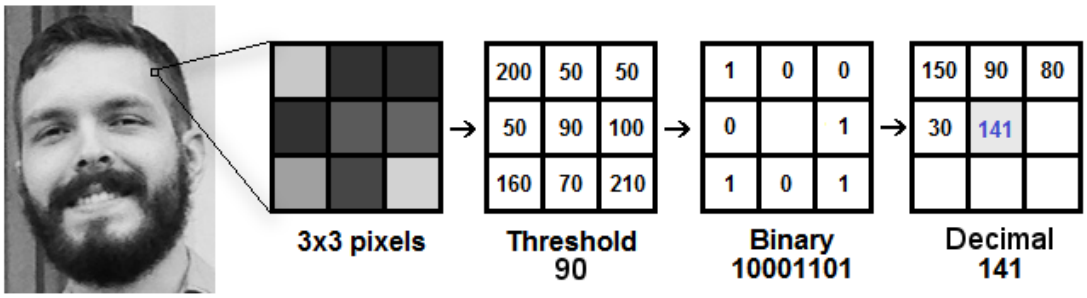
Paulo Viola e Michael Jones escreveram um artigo em 2001 intitulado *Rapid Object Detection using a Boosted Cascade of Simple Features*, nesse artigo, é defendida a ideia de um método eficaz de detecção rápida de objeto usando classificador em cascata baseado no recurso Haar. Essa é uma abordagem de aprendizado de máquina onde a função cascata é treinada a partir de muitas imagens positivas e negativas.

Quando estamos falando em detecção facial é entendido que esse aprendizado é dado a partir de imagens positivas, que tem faces, e imagens negativas, que não tem faces. A partir desse treinamento obtemos o classificador de face. A grande vantagem desse tipo de recurso é a velocidade que ele identifica o objeto de interesse. Uma vez detectada uma face é dada sequência em algum outro algoritmo para a identificação facial, a quem ela pertence.

3.6 LBPH

LBPH é um combinado de duas técnicas, *Local Binary Patterns* (LPB) e *Histograms of Oriented Gradients* (HOG). LBP é um tratamento a partir da imagem a ser analisada onde é extraído uma matrix 3x3 dos pixels e são analisados os pixels no entorno do pixel central da matriz e, caso o valor do pixel seja maior que o valor do pixel central, ele é convertido em um (1), caso seja menor é convertido em zero (0), conforme ilustra a Figura 2. É importante ressaltar que esse tratamento acontece apenas em imagens convertidas para tons de cinza, tons esses que variam em intensidade de 0 a 255.

Figura 2 – Tratamento do algoritmo LBP

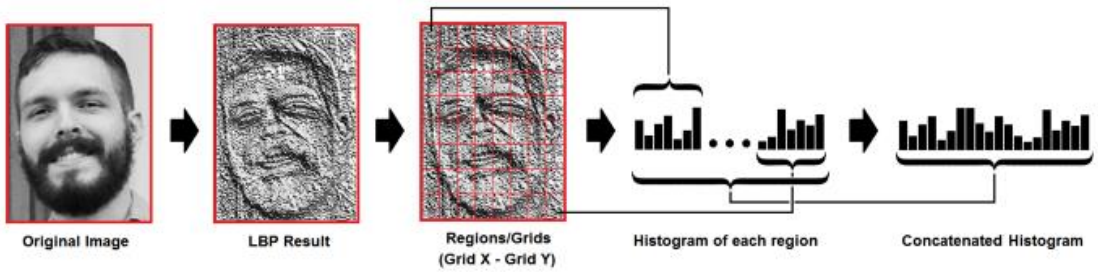


Fonte: Imagem do site Towards Data Science[2]

Como é ilustrado na Figura 2, o resultado obtido é uma matriz de zeros e uns. Os binários são colocados em sequência, representando um tipo de 8 bits, que são convertidos para decimal. Esse valor decimal é atribuído ao pixel do meio da matriz 3x3. Esse processo é executado para todos os pixels das imagens, gerando uma nova imagem (TOWARDS DATA SCIENCE).

Após a execução desse tratamento, a imagem resultante é convertida em histograma. Nele são dispostos todos os pixels do resultado do tratamento LBP. Como a intensidade dos pixels varia de 0 a 255, obtemos um histograma de 256 colunas. O histograma é quem carrega a identificação fácil, ele é o resultado final das combinações das técnicas LBP e HOG, chamada de algoritmo LBPH. A Figura 3 ilustra o histograma resultante.

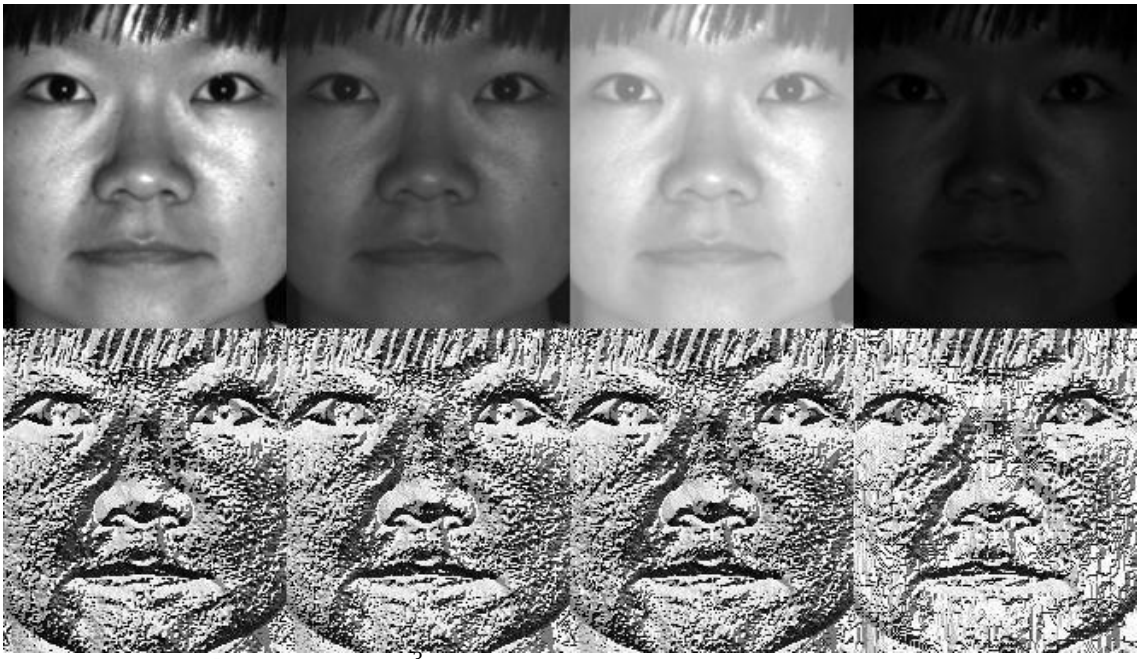
Figura 3 – O emprego do HOG e o resultado do LPBH



Fonte: Imagem do site Towards Data Science[2]

Uma característica dessa técnica é que ela consegue destacar o relevo das imagens das faces, outra característica que é destaque dessa técnica é a robustez para variações de luminosidade. A Figura 4 exemplifica bem as variações de luminosidade e o resultado de tratamento LBP para as distintas intensidades de luminosidades, onde é possível notar muita pouca diferença.

Figura 4 – Face com diferentes intensidades de iluminação com os respectivos resultados do tratamento do algoritmo LBP



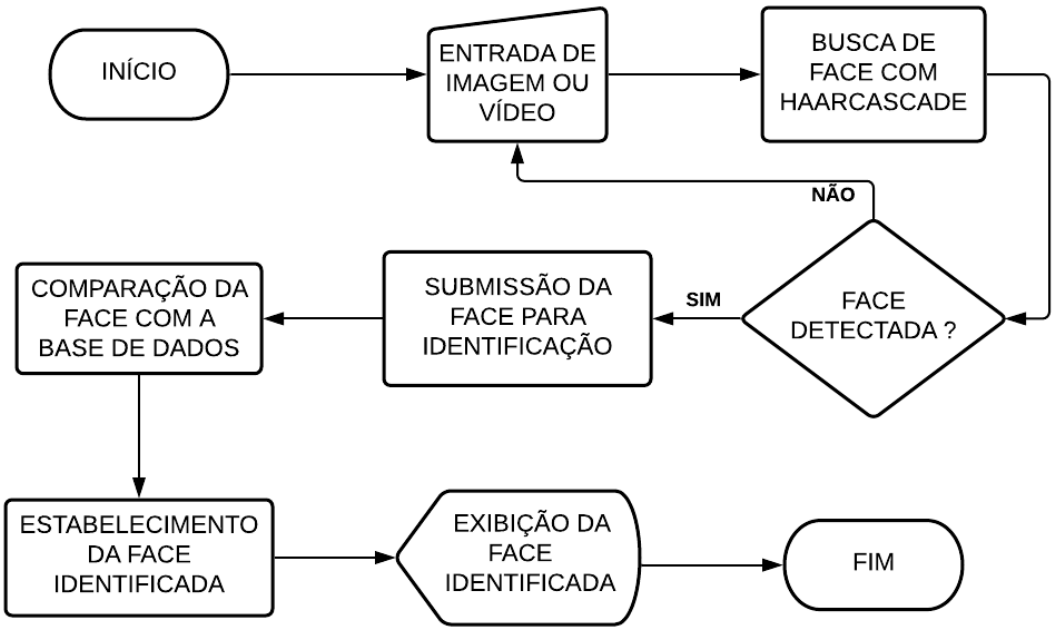
Fonte: Imagem do site OpenCV[3]

Após executado esse tratamento é feito um histograma com todos dos decimais resultantes. A identificação facial acontece a partir da comparação com os histogramas da base de dados.

4. FLUXOGRAMA

O fluxograma da Figura 5 demonstra as etapas que o sistema proposto executa para a realização da detecção e identificação facial. Nele ocorre primeiro a entrada dos dados, que são as imagens onde queremos detectar e identificar as faces. Caso seja reconhecida a face ela é submetida a identificação, sendo realizada a comparação das faces reconhecidas com as faces do banco de dados. Feito a identificação é mostrado ao usuário a imagem inserida no sistema com as faces marcadas com uma borda quadrada no seu entorno e o nome da pessoa a qual ela pertence logo abaixo da borda quadrada.

Figura 5 – Imagem do fluxograma mostrando o processo para a detecção e identificação facial no projeto proposto



Fonte: Próprio do autor.

5. CONSTRUÇÃO DO SISTEMA

Para a construção do sistema será usado o Raspberry Pi 3 Model B+. Nele foi instalado a biblioteca OpenCV versão 4.1.0 pois, como já dito, é nessa biblioteca que se encontram os algoritmos de visão computacional para a detecção e identificação facial. Até a data de produção desse trabalho já foram disponibilizadas versões do OpenCV mais atualizadas, entretanto, para o objetivo desse trabalho não é exigido a versão mais recente, poderíamos até mesmo usar versões mais antigas que também conseguiríamos desenvolver.

Como câmera foi usado a câmera do smartphone pois, através do uso de um aplicativo chamado DroidCam, é possível usar a câmera de qualquer smartphone com sistema operacional Android como webcam, ou mesmo como câmera IP. O smartphone estando conectado na mesma rede que o dispositivo embarcado, nesse caso o Raspberry Pi, o fluxo de dados capturados pela câmera é compartilhado na rede em que ambos os dispositivos estejam conectados. Ou ainda, podemos usar um cabo USB para conectar os dispositivos e assim conseguir orientar o fluxo de dados capturados pela câmera do smartphone.

Na IDLE foram construídas três aplicações. A primeira aplicação tem como objetivo a construção da base de dados, isso é, foi ela quem capturou as imagens dos candidatos os quais queremos que o sistema identifique. Esses dados capturados, que são imagens, foram usados para o treinamento do algoritmo de inteligência artificial. A segunda aplicação é o treinamento do algoritmo. Como já explanado, o algoritmo LBPH extrai características de cada uma das imagens. Como resultado obtemos um arquivo .yml e é nesse arquivo que consta as características de cada uma das pessoas da base de dados. A terceira aplicação será a detecção e identificação em si, onde a câmera passa a captura de imagens do ambiente. Durante essa captura o HaarCascade procurar faces em cada um dos frames capturados. Uma vez encontrada a face, a aplicação marcar essa face com bordas, formando um quadrado no entorno da face detectada. Em seguida, a aplicação faz a comparação dessa face com as faces do arquivo .yml e caso a face seja pertencente a alguém da base de dados, a aplicação escrever o nome da pessoa identificada abaixo da borda quadrada. É esperado com isso a detecção e identificação facial em tempo real.

6. RESULTADOS E DISCUSSÃO

A construção do sistema foi realizada com êxito, com algumas ressalvas. Das aplicações, a primeira, a etapa de captura da face para abastecer a base de dados, funcionou sem qualquer problema. Essa é uma etapa fundamental para o funcionamento do sistema uma vez que é dela que é gerado o arquivo .yml, resultante do treinamento do algoritmo. Para a extração das faces na primeira aplicação, foi usado um vídeo para a extração das faces cada uma das pessoas que compos a base de dados. Cada pessoa fez um vídeo de si mesma, com pouco menos de 1 minuto e cada vídeo rendeu mais de mil faces. Embora a quantidade de faces seja grande, a base de dados continha apenas quatro pessoas. Não há necessidade de ser esse número tão grande de faces para o treinamento do algoritmo. 10 imagens já é o suficiente para o funcionamento do sistema.

Na segunda aplicação, que é o treinamento do algoritmo em si, houve um imprevisto. Durante o treinamento do algoritmo, com essa grande quantidade de faces, o Raspberry Pi sofreu um *overflow* na memória, interrompendo o treinamento, lembrando que o modelo do Raspberry Pi usado tem 1Gb de memória RAM. Como alternativa, foi feito esse treinamento numa máquina com mais memória e maior poder de processamento e, uma vez o algoritmo treinado, foi copiado o arquivo .yml resultante do treinamento no Raspberry Pi.

A terceira aplicação, onde de fato ocorre a detecção e identificação facial, apresentou algumas limitações conforme os resultados a seguir.

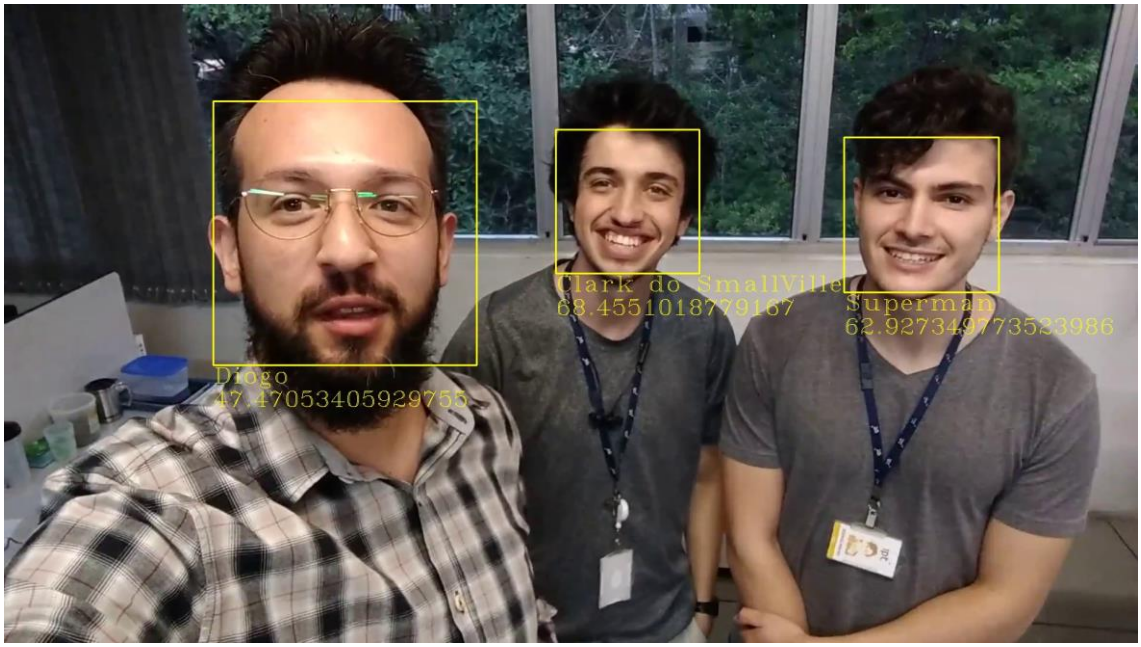
6.1 Resultados

As Figuras 6, 7 e 8 apresentam algumas imagens geradas por esse sistema, mostrando as possibilidades de implantação desse algoritmo em projetos para detecção e identificação facial. As imagens apresentadas são: frontal, lateral, com sorriso, sério. A identificação foi verificada em todas as imagens.

Essas imagens exibem um quadrado no entorno da face detectada e logo abaixo do quadrado é exibido os dados da identificação facial, caracterizado pelo nome da pessoa e um parâmetro relacionado comparação com a base de dados.

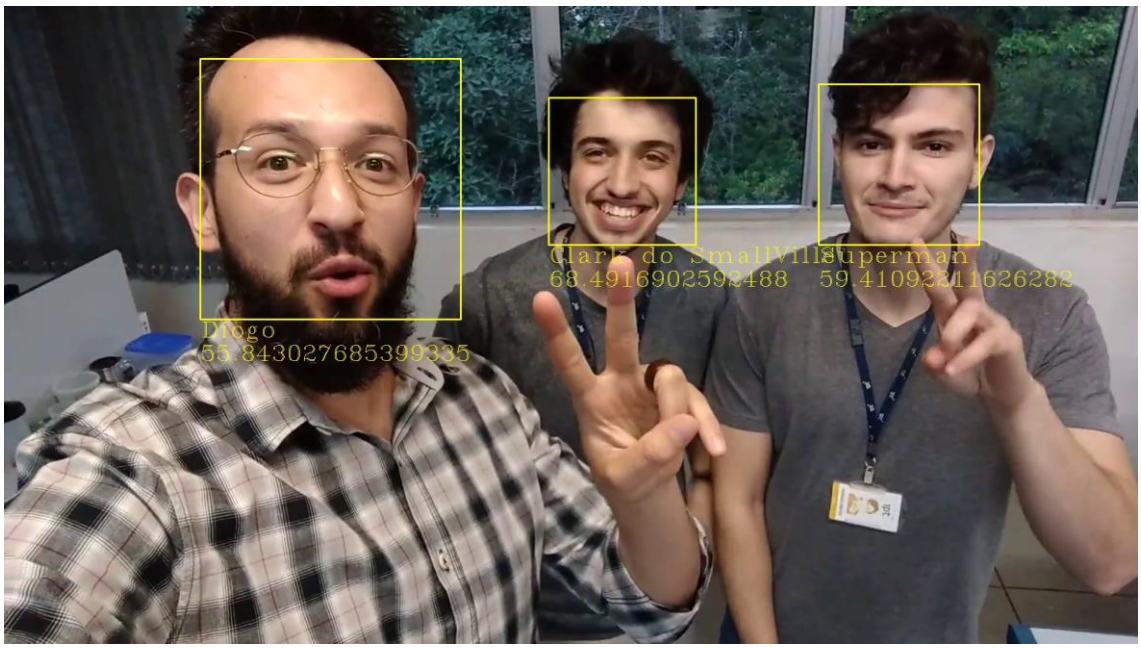
Figura 6 – Imagens capturadas pela câmera e a detecção e identificação das pessoas.

a) Imagem frontal



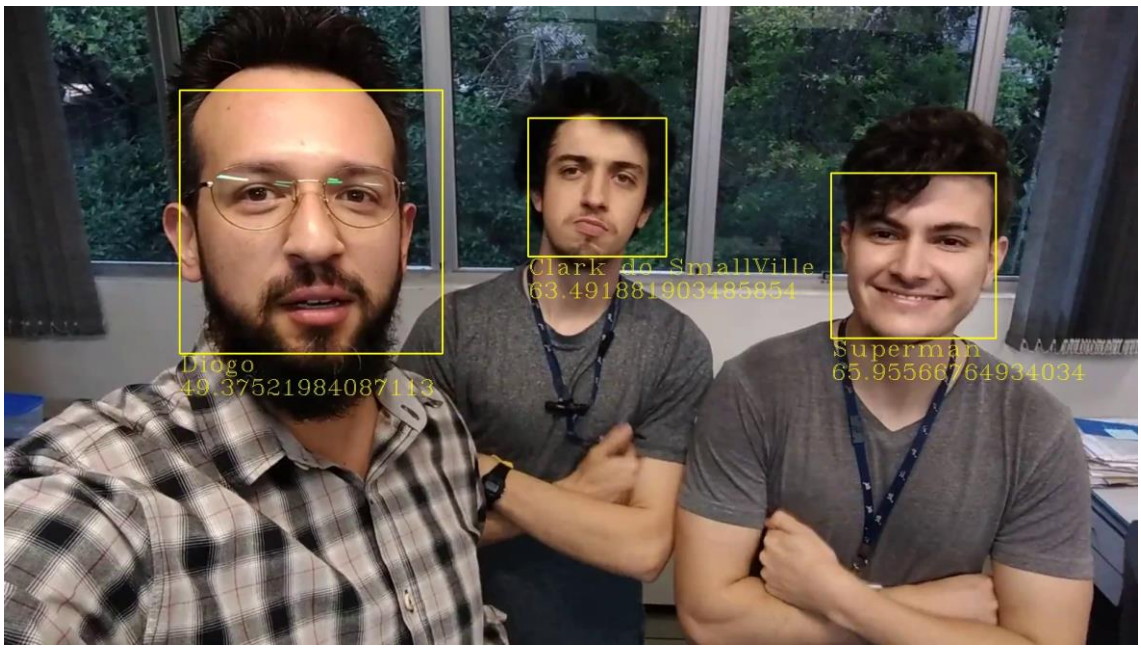
Fonte: Próprio do autor

b) Imagem lateral



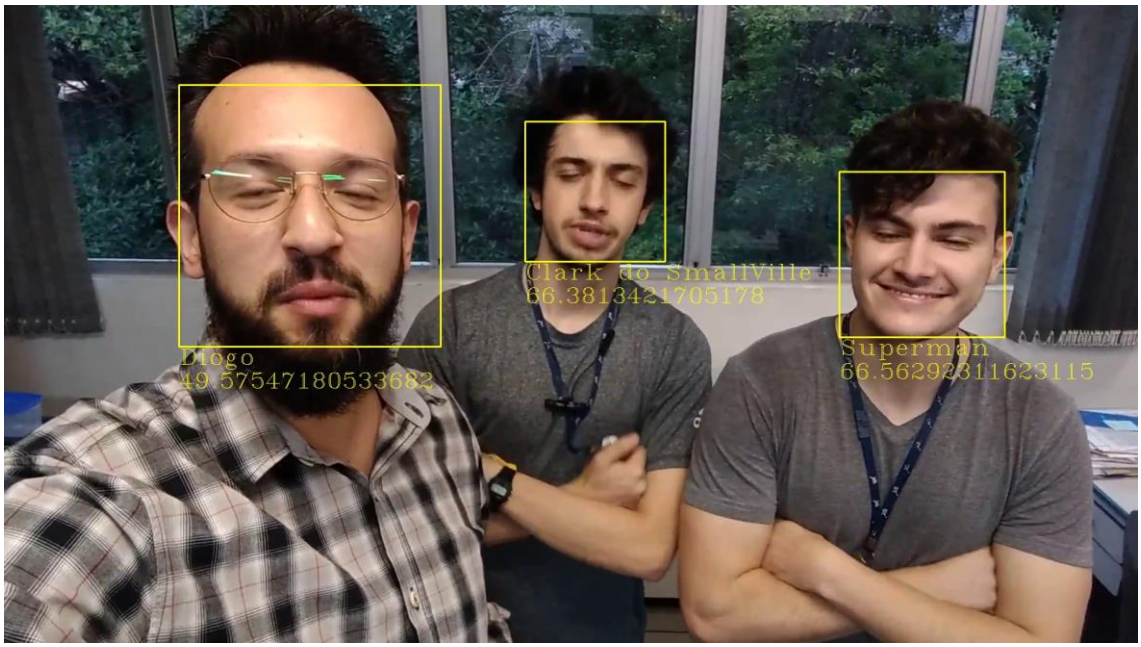
Fonte: Próprio do autor

c) Imagem com sorriso



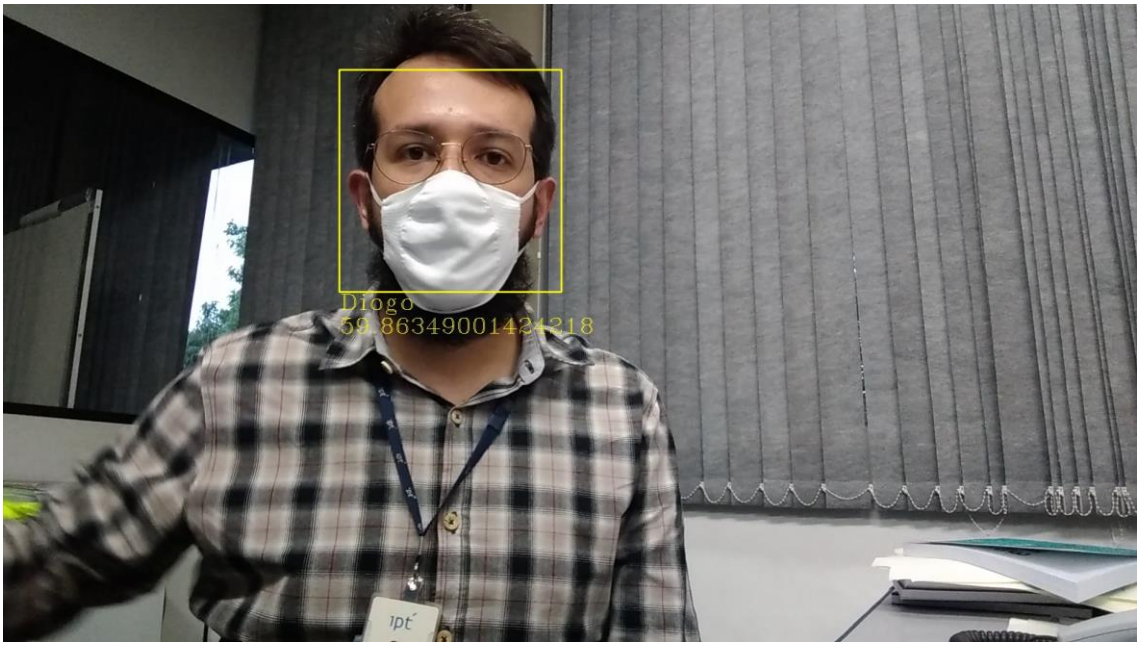
Fonte: Próprio do autor

d) Imagem com olhos fechados



Fonte: Própria do autor

e) Imagem rosto com máscara



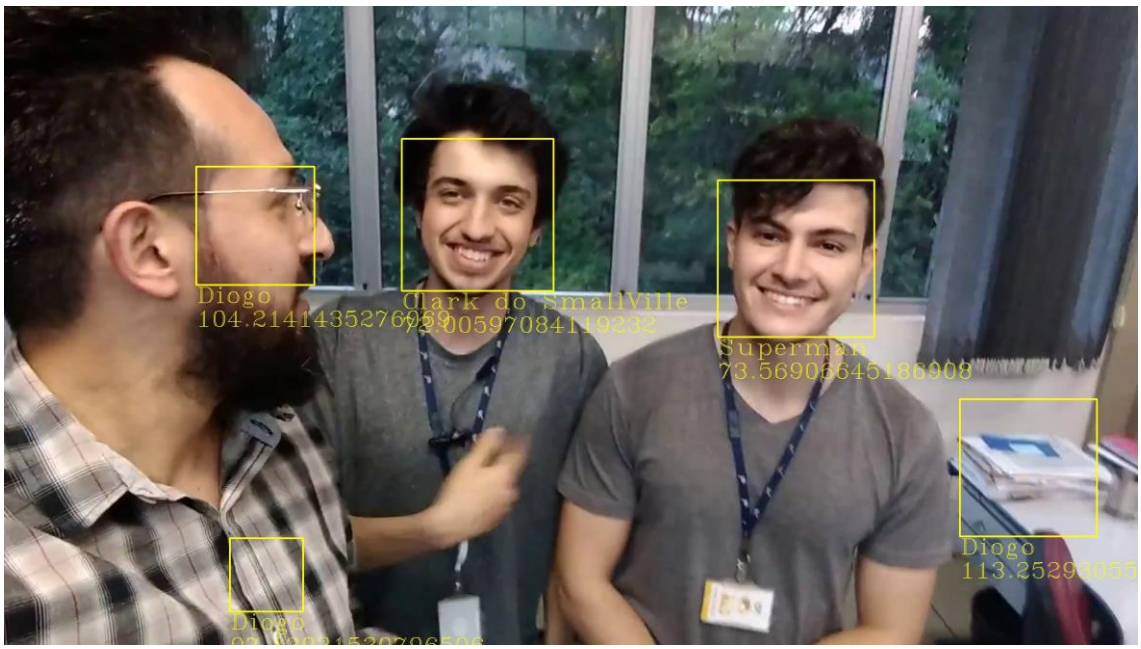
Fonte: Própria do autor

A Figura 6 ilustra bem a detecção e identificação facial em diferentes concisões faciais, inclusive quando a pessoa está usando máscara, mostrando que o emprego dessa técnica funciona bem, tornando-a passível de ser aplicada em diversas situações.

A Figura 7 mostra um dos problemas dessa técnica, o falso positivo. Como já posto, antes da identificação facial em si, primeiro ocorre uma detecção facial. A técnica empregada para a detecção é do HaarCascade. Ocorre que, o HaarCascade pode detectar faces na imagem que não são faces, são falsos positivos. Nesse sistema foi usado uma distribuição popular do HaarCascade para detecção de faces. Há muitos fóruns na internet que recomendam fortemente que o desenvolvedor do sistema treine o próprio HaarCascade para o emprego de detecção de faces, pois algumas distribuições públicas costumam gerar esses falsos positivos.

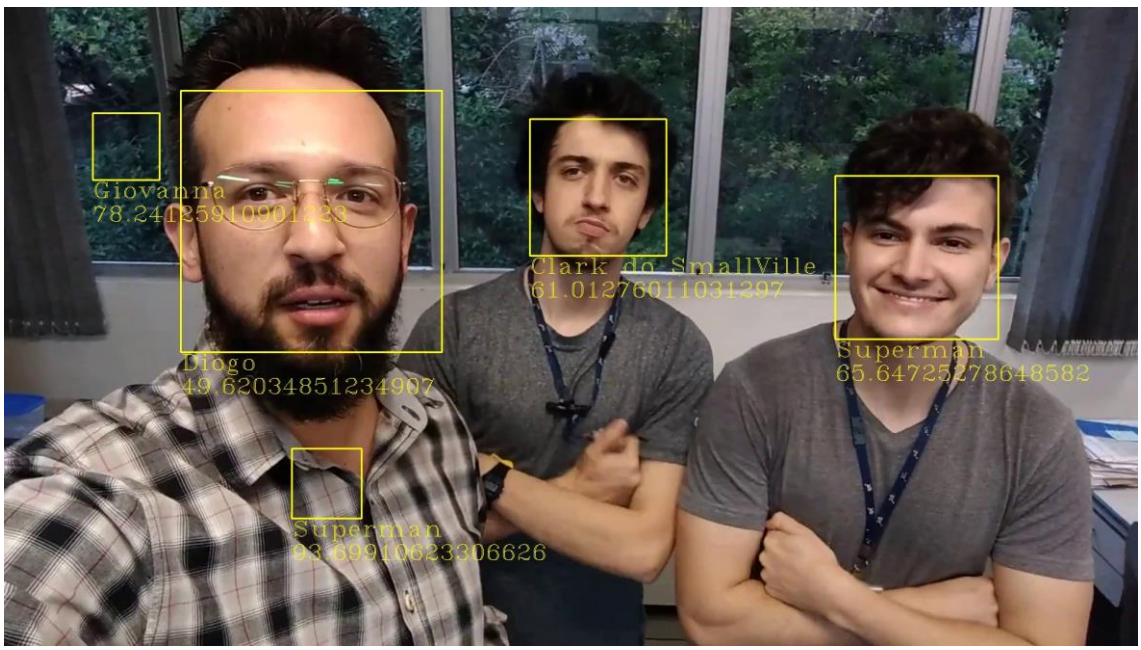
Figura 7 – Imagens com detecções de faces falsas.

a) Imagens com falsos positivos



Fonte: Própria do autor

b) Imagens com falsos positivos

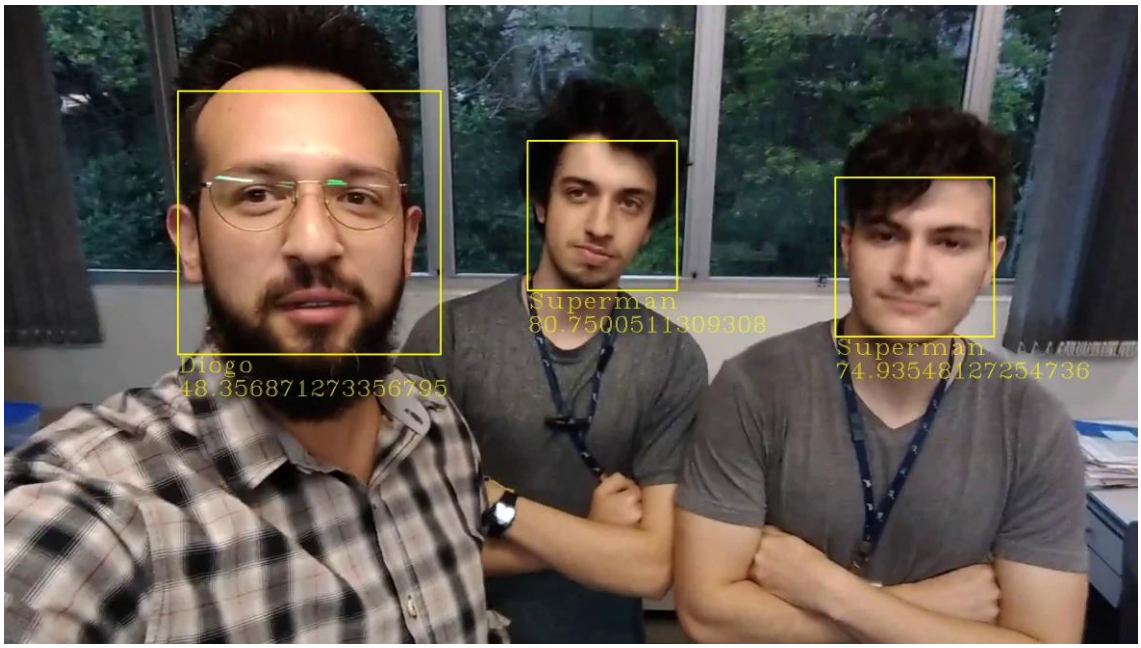


Fonte: Própria do autor

Na Figura 8 é possível notar que a identificação facial ocorreu com falha, todas as faces foram detectadas na imagem, entretanto a identificação facial não correspondeu conforme o esperado. É notado na imagem que a identificação facial reconheceu a mesma pessoa em duas faces diferentes.

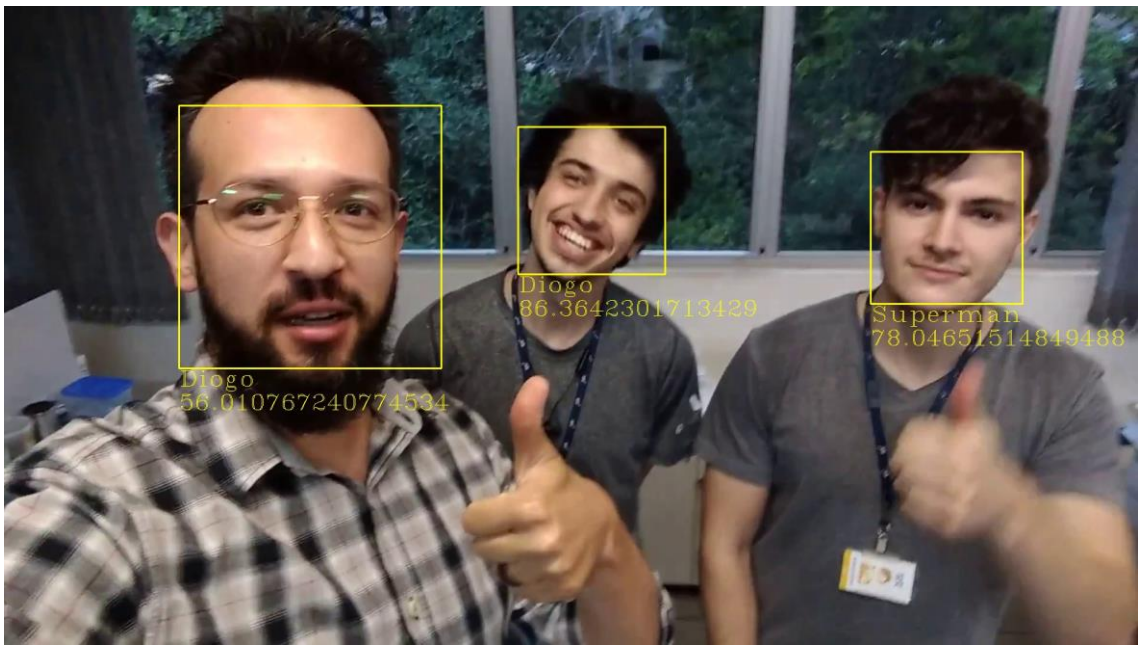
Figura 8 – Imagens com falha na identificação de uma face.

a) Imagem identificação errada



Fonte: Própria do autor

b) Imagem identificação errada

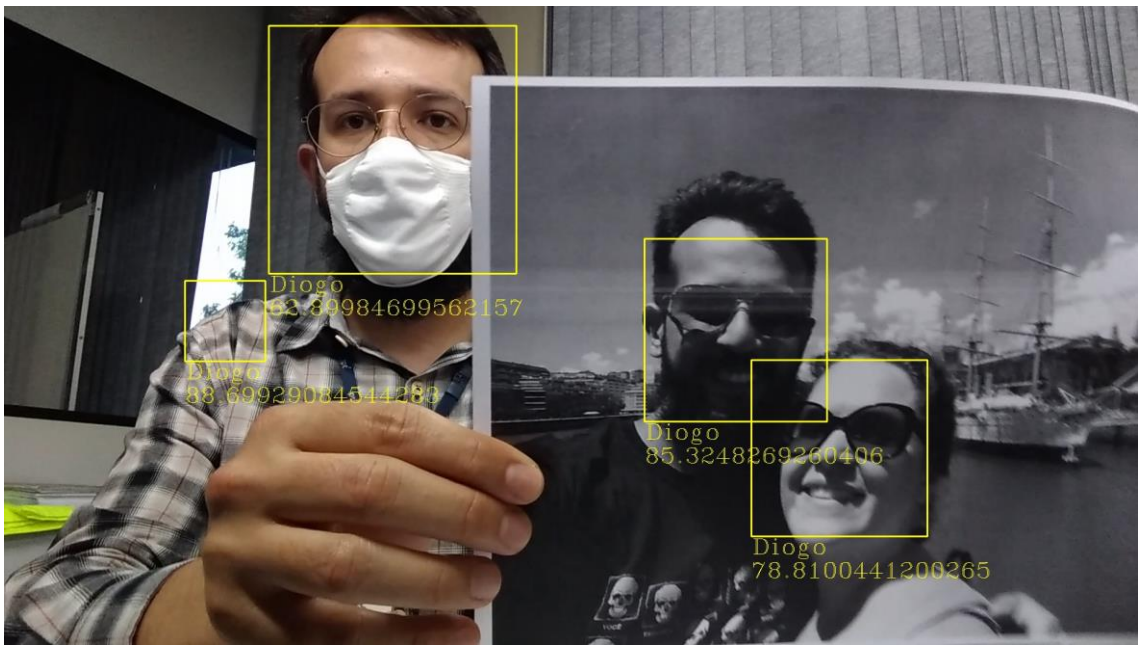


Fonte: Própria do autor

A Figura 9 retrata uma questão que não é desejada. Foi detectada face em uma foto, o que mostra uma vulnerabilidade do sistema caso ele seja usado para validação de biometria facial, mesmo que a identificação não tenha sido correta.

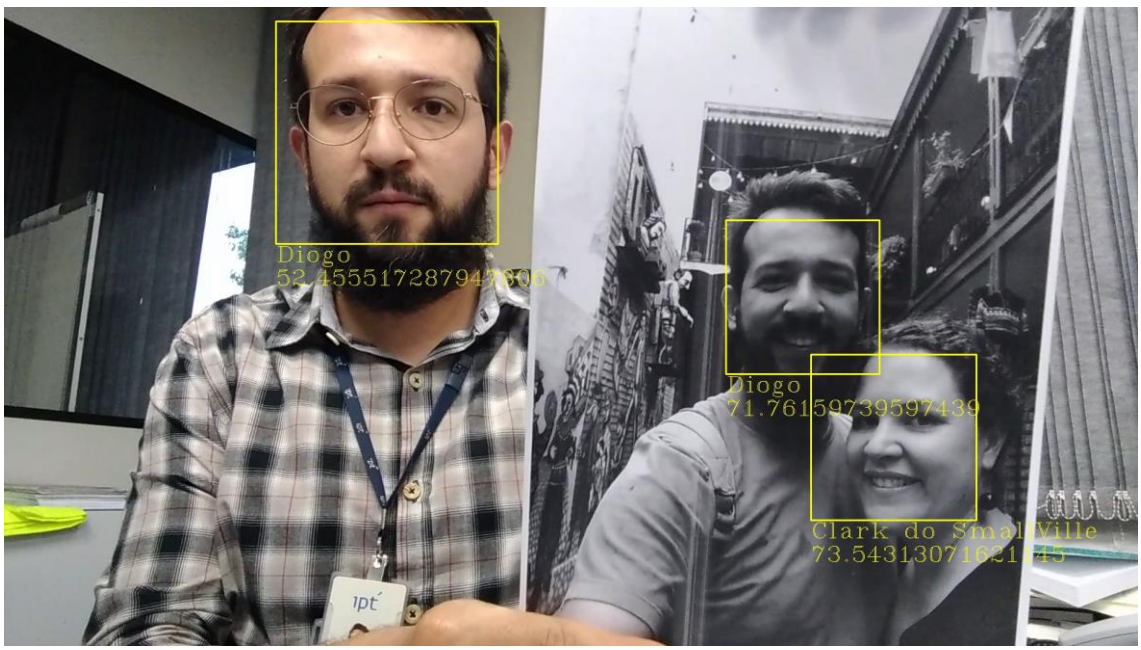
Figura 9 – Imagens de faces detectadas em fotos

a) Imagem de faces detectadas em fotos, identificação errada e falso positivo.



Fonte: Própria do autor

b) Imagem de faces detectadas em foto e identificação errada



Fonte: Própria do autor

6.2 Limitações

Uma primeira limitação está no tempo de transmissão dos dados das imagens capturadas. O fato de a imagem ser capturada por um smartphone e precisar ser modulada para trafegar na rede, resulta que, ao chegar no dispositivo em que será processada, ele sofrerá um outro processo para demodulação e isso em si gera um delay. A qualidade do dispositivo transmissor dos dados, a rede que o dado vai trafegar e a qualidade do dispositivo receptor dos dados também influenciam nos tempos de resposta. Nas condições que esse sistema funcionou, o tempo de captura das imagens até o processamento e exibição do resultado gerou uma demora de aproximadamente 4 segundos. Podemos interpretar esse tempo como um grande atraso, mas, ainda sim, plausível, a depender do emprego desse tipo de aplicação.

Outra questão importante que é claramente perceptível são os falsos positivos. O HaarCascade confunde alguns trechos das imagens que estão sendo tratadas como uma face, quando na verdade são outras coisas, não faces propriamente. Além disso, uma vez que o HaarCascade captura essa “falsa face”, ela é tratada como uma face real e é comparada à base de dados contida no arquivo .yml gerado. Finalmente a “falsa face” é identificada como alguém da base de dados, configurando uma situação indesejada.

Isso representa uma limitação importante haja vista que o computador faz uma comparação métrica entre a face capturada e a base de dados para a identificação facial. Caso a face capturada não esteja na base de dados, essa face fatalmente receberá a identificação de algum indivíduo da base de dados. Esse é um problema que ainda é possível ser contornado estipulando um valor melhor para o parâmetro que compara a menor distância entre a face capturada e as faces da base de dados.

Nesse sistema não é possível distinguir uma pessoa real capturada pelo HaarCascade de um retrato ou uma foto, por exemplo. Isso representa um perigo caso esse tipo de sistema seja para uso de autenticação biométrica, pois, qualquer um que use a foto da pessoa a ser identificada na frente do próprio rosto inevitavelmente a face da foto será detectada e identificada.

6.3 Potenciais Aplicações

Tendo em vista as limitações desse tipo de sistema e pensando exclusivamente na aplicação do algoritmo LBPH para a detecção e identificação facial, há muitas possibilidades de emprego desse tipo de sistema no mercado.

Num sistema de caixa eletrônico em bancos, essa aplicação pode, e em muitos bancos já é, ser usado como um segundo ou terceiro elemento de autenticação para prevenção contra possíveis fraudes como saques de dinheiro da conta por terceiros ou furto de cartões, pois, a face humana, assim como a digital, é considerada como identificação biométrica.

As empresas que usam crachá para registrar o ponto de funcionários, com o uso de apenas uma câmera, conseguirão reconhecer todos os trabalhadores que chegam registrando presença e hora dessa chegada.

Muitos aplicativos de redes sociais já empregam massivamente o uso de detecção facial para filtro ou mesmo para sugerir agrupamentos de imagens em determinadas categorias.

Ambiente industriais restritos como laboratórios químicos ou salas de ambiente controlado, de segurança, podem ser acessados com o uso de detecção e identificação facial usando autenticação de biometria facial.

Muitas aplicações podem ser pensadas para o emprego da técnica de detecção e identificação facial, entretanto, ainda há um número pequeno de profissionais que dominam as técnicas de visão computacional para empregar mais aplicações nos distintos setores passíveis do seu uso.

7. CONCLUSÃO

A proposta de construção de um de sistema de detecção e identificação facial embarcado em dispositivo de baixo custo aconteceu conforme foi ilustrado em tópicos anteriores, entretanto, melhorias podem ser empregadas para aumentar a robustez do sistema como um todo.

8. SUGESTÃO DE MELHORIA

Uma sugestão para o melhoramento do sistema quanto a velocidade de transmissão é garantir que o fluxo dos dados ocorra exclusivamente por cabo. Nas câmeras IPs que têm os dados trafegados por cabo o processamento dos dados ocorre de forma muito mais rápida, diferente dos dados que trafegam por sinal WiFi, que exigem uma modulação e demodulação do sinal, gerando não apenas um delay na transmissão dos dados como também uma perda da qualidade do sinal.

A manipulação dos parâmetros que envolvem a técnica do algoritmo LBPH também deve ser explorada para o melhoramento do sistema no seu todo, por exemplo, a alteração nos parâmetros mínimo tamanho da face a ser detectada.

O hardware que roda o sistema é uma limitação importante, pois, as tecnologias que envolvem tratamento de imagem exigem maior poder de processamento.

Agora, se queremos uma melhor resposta para a identificação facial, sem dúvida que devemos apontar como proposta o uso das Redes Neurais. Alguns estudos citados nesse trabalho (SILLES, 2020) e inúmeros trabalhos já realizados em muitas academias e institutos mostram que as Redes Neurais de longe oferecem uma qualidade muito superior frente às demais técnicas de Visão Computacional. Tamanha robustez de resultado ofertado pelas Redes Neurais exige maior poder de processamento. Felizmente, hoje existem hardwares especializados em rodar Redes Neurais a fim de facilitar e tornar mais eficiente o emprego dessa técnica nas mais distintas aplicações potenciais a serem desenvolvidas.

REFERÊNCIAS

BARELLI, Felipe. **Introdução à Visão Computacional: Uma abordagem prática com Python e OpenCV**. Rio de Janeiro: Editora Casa do Código, 2018.

DEBIAN PROJECT. **Welcome to Raspbian.** Página inicial. Disponível em: https://www.raspbian.org/FrontPage. Acesso em: 10 de set. de 2020.

[3]FACE RECOGNITION WITH OPENCV. **Local Binary Patterns Histrograms.** Disponível em: <https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html#local-binary-patterns-histograms>. Acessado em: 15 de set. de 2020.

LEE, Kai-Fu. **Inteligência artificial: como os robôs estão mudando o mundo, a forma como amamos, nos comunicamos e vivemos**. Tradução Marcelo Brabão. Rio de Janeiro: Globo Livros, 2019.

MAGRANI, Eduardo. **A internet das coisas**. Rio de Janeiro: FGV Editora, 2018.

MANSUR, Ricardo. **Inteligência Estendida com Inteligência Artificial**. Rio de Janeiro: Editora Ciência Moderna Ltda., 2019.

MERCES, Ricardo**. Raspberry Pi – Conceito & Prática**. 2ª Edição. Rio de Janeiro: Editora Ciência Moderna Ltda. 2014.

OLIVEIRA, Adriano. O que você deve saber antes de adquirir uma câmera térmica para combater a Covid-19. Revista Segurança Eletrônica. Osasco. Ano 4. Nº 41. Junho de 2020. Pág 42.

OPENCV TEAM. **About.** Disponível em: <https://opencv.org/about/>. Acessado em: 11 de set. de 2020.

PYTHON SOFTWARE FUNDATION. **IDLE.** Disponível em: < https://docs.python.org/3/library/idle.html>. Acessado em: 10 de set. de 2020.

PYTHON.ORG. **BeginnersGuide.** Disponível em: <https://wiki.python.org/moin/BeginnersGuide/Overview>. Acessado em: 14 de set. de 2020.

[1]RASPBERRY PI. Raspberry Pi 3 Model B+. Disponível em: <https://www.raspberrypi.org/products/raspberry-pi-3-model-b-plus/?resellerType=home>. Acessado em: 18 de nov. de 2020.

REVISTA DIGITAL SECURITY. ARTIGO: A evolução das câmeras IP e sua influência no mercado de segurança eletrônica. Disponível em <https://revistadigitalsecurity.com.br/artigo-a-evolucao-das-cameras-ip-e-sua-influencia-no-mercado-de-seguranca-eletronica/>. Acessado em: 10 de dez. de 2020.

REVISTA SEGURANÇA ELETRÔNICA. **10 fatos interessantes sobre CFTV.** Disponível em <https://revistasegurancaeletronica.com.br/10-fatos-interessantes-sobre-cftv/>. Acessado em: 10 de dez. de 2020.

SILLES, Felipe Silva. **Reconhecimento Facial: Seleção de técnicas para processamento em dispositivos compactos**. Dissertação (Mestrado) — Instituto de Pesquisas Tecnológicas do Estado de São Paulo S.A., São Paulo, 2020.

SILVA NETO, Antônio José da. **Inteligência computacional aplicada a problemas inversos em transferência radiativa**. Rio de Janeiro: EdUERJ, 2016.

[2]TOWARDS DATA SCIENCE. **Face Recognition: Unerstanding LBPH Algorithm.** Disponível em: <https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b>. Acessado em: 11 de nov. de 2020.