Universidade de Brasília - UnB
Faculdade UnB Gama - FGA
Engenharia Eletrônica

Estimação de frequência cardíaca e saturação de oxigênio a partir de segmentos de vídeo usando técnicas de processamento de imagem

Autor: Rodrigo Ribeiro Lopes Trindade
Orientador: Prof. Dr. Marcus Vinícius Chaffim Costa

Brasília, DF
2024



Rodrigo Ribeiro Lopes Trindade

Estimação de frequência cardíaca e saturação de oxigênio a partir de segmentos de vídeo usando técnicas de processamento de imagem

Monografia submetida ao curso de graduação em Engenharia Eletrônica da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia Eletrônica.

Universidade de Brasília - UnB

Faculdade UnB Gama - FGA

Orientador: Prof. Dr. Marcus Vinícius Chaffim Costa

Brasília, DF

2024

Rodrigo Ribeiro Lopes Trindade

Estimação de frequência cardíaca e saturação de oxigênio a partir de segmentos de vídeo usando técnicas de processamento de imagem/ Rodrigo Ribeiro Lopes Trindade. – Brasília, DF, 2024-

72 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Marcus Vinícius Chaffim Costa

Trabalho de Conclusão de Curso – Universidade de Brasília - UnB
Faculdade UnB Gama - FGA , 2024.

1. Estimação da Frequência Cardíaca e Saturação de Oxigênio. 2. Processamento de imagem. I. Prof. Dr. Marcus Vinícius Chaffim Costa . II. Universidade de Brasília. III. Faculdade UnB Gama. IV. Estimação de frequência cardíaca e saturação de oxigênio a partir de segmentos de vídeo usando técnicas de processamento de imagem

CDU 02:141:005.6

Rodrigo Ribeiro Lopes Trindade

Estimação de frequência cardíaca e saturação de oxigênio a partir de segmentos de vídeo usando técnicas de processamento de imagem

Monografia submetida ao curso de graduação em Engenharia Eletrônica da Universidade de Brasília, como requisito parcial para obtenção do Título de Bacharel em Engenharia Eletrônica.

Trabalho aprovado. Brasília, DF, Setembro de 2024:

Prof. Dr. Marcus Vinícius Chaffim Costa
Orientador

Prof. Dr. Diogo Caetano Garcia
FGA/UnB
Convidado 1

Prof. Dr. Flávio de Barros Vidal
CIC/UnB
Convidado 2

Brasília, DF
2024

Agradecimentos

Gostaria de iniciar agradecendo a Deus, pelas oportunidades que Ele me trouxe, pelas pessoas que colocou em minha vida, que foram cruciais para o meu desenvolvimento e sobrevivência, e pelo Seu amor incondicional que sempre restaurou minhas forças, iluminou minhas trevas e guiou-me aos caminhos da esperança. A Ele agradeço a todas as coisas boas em minha vida.

Gostaria de agradecer a minha família que sempre me incentivou, apoiou, cobrou e se fez presente, para que eu nunca desanimasse e sempre encontrasse forças e motivos para lutar, por eles que nunca desistiram de mim, mesmo nos momentos mais difíceis, e muito pelo contrário, sempre foram restauração a minha alma. Agradeço aos meus pais por me darem a chama para lutar, ao meu irmão por me inspirar em tudo que sou e a minha irmã por renovar minha vida, me ensinar novas coisas e me inspirar com sua fantástica inteligência. Devo a eles tudo o que a de bom e bem sucedido em mim e em minha vida.

Agradeço também a minha amada namorada e fiel companheira que desde o dia 05 de Dezembro de 2020 se fez presente em cada dia de minha vida, sempre me alegrando com seu sorriso e suas risadas fofas, que sempre me acompanhou em cada etapa da minha jornada, sempre me deu esperanças e sempre me motivou. Agradeço a esta princesa que, ainda que as vezes não percebesse, sempre curou minhas feridas. A ela devo meu coração e minha alegria.

Agradeço ao professor Marcus Chaffim, meu orientador do TCC, e ao professor Daniel Araújo, meu orientador de pesquisa, que me ensinaram bastante nesta minha jornada profissional e não pouparam esforços para me ajudar, tanto como aluno, quanto como pessoa. Estes não se limitaram a me ensinar, com muita excelência por sinal, sobre assuntos da engenharia, mas também sobre a vida humana. Agradeço também aos professores Ricardo Fragelli, Ronni Amorim, Vinícius Rispolli e Suélia Fleury Rosa, cujos nome menciono na ordem em que conheci. Vocês me inspiraram em diversos detalhes tanto profissionalmente, quanto como humanos. Cade um dos seis mencionados me proporcionaram momentos incríveis que renderam um afeto e admiração que nunca se acabará.

Por fim, não poderia deixar de agradecer aos meus companheiros de guerra, meus amigos e colegas da graduação. Cada momento de parceria, raiva e alegria, cada risada, cada conhecimento compartilhado estará eternamente guardado em meu coração. Destes dou destaque aos irmãos que adquiri durante minha vida, Leoni Gabriel e Lucas Pereira, que me acompanham a uma longa jornada. Obrigado pelo carinho, apoio e principalmente paciência. Vocês são minha história.

Resumo

O cenário pandêmico da Covid-19, vivenciado em períodos próximos à escrita deste documento, evidenciou a necessidade de realizar determinadas tarefas de forma remota em ocasiões específicas. Além disso, a busca por métodos não-invasivos e sem contato físico com pacientes é bastante promissora no mundo contemporâneo, pois permite uma análise mais prolongada e profunda sem causar desconforto ao paciente. Fundamentado nessas duas realidades, este trabalho apresenta uma proposta da detecção da frequência cardíaca e da saturação de oxigênio utilizando segmentos de vídeo e programação em *Python*. As imagens utilizadas para comprovar a eficácia do sistema proposto e para ajuste do mesmo foram obtidas através das câmeras do celular *Samsung Galaxy 30s* e *POCO X4 Pro 5G*, e por meio do vídeo disponibilizado por Wu *et al.* e foram processadas em um computador com processador I7 da sétima geração. O projeto se divide em três etapas, sendo elas: detecção facial e rastreamento dos pontos desejados, tratamento dos quadros de vídeo e cálculo dos parâmetros desejados. Os resultados foram eficientes, apresentando robustez a diferentes tipos de iluminação e taxas de transmissão, porém com limitações quanto a mudanças de usuários e variações bruscas de luz.

Palavras-chaves: Landmarks, Detecção facial, Python, Saturação de oxigênio, Frequência Cardíaca, EVM, RGB, YIQ, YCbCr, Lab.

Abstract

The Covid-19 pandemic scenario, experienced in periods close to the writing of this document, highlighted the need to perform certain tasks remotely in specific situations. Additionally, the search for non-invasive methods and without physical contact with patients is quite promising in the contemporary world, as it allows for a more prolonged and in-depth analysis without causing discomfort to the patient. Based on these two realities, this work presents a proposal for detecting heart rate and oxygen saturation using video segments and programming in Python. The images used to prove the effectiveness of the proposed system and to adjust it were obtained through the cameras of the Samsung Galaxy 30s and POCO X4 Pro 5G cell phones, and the video provided by Wu et al., and were processed on a computer with a seventh-generation I7 processor. The project is divided into three stages: facial detection and tracking of the desired points, video frame processing, and calculation of the desired parameters. The results were efficient, showing robustness to different types of lighting and transmission rates, but with limitations regarding user changes and abrupt light variations.

Key-words: Landmarks, Facial Detection, Python, RGB, YIQ, YCbCr, Lab, Oxygen Saturation, Heart Rate.

Lista de ilustrações

Figura 1 – Morfologia interna do coração 27
Figura 2 – Sistema de Condução Elétrica do Coração 29
Figura 3 – Sinal ECG registrado posicionado ao lado do coração.
Fonte: Adaptado de (YITZHAKNAT, 2022) 31
Figura 4 – Definições das ondas do ciclo cardíaco e suas durações. 32
Figura 5 – Onda da fotopletismografia 34
Figura 6 – Ondas do sinal PPG. 34
Figura 7 – Detecção Facial com *Landmarks* Selecionados Destacados 36
Figura 8 – Gráfico comparativo entre os valores estimados e os valores reais para a frequência cardíaca. 47
Figura 9 – Gráfico comparativo entre os valores estimados e os valores reais para a saturação de oxigênio. 47
Figura 10 – Detecção facial realizada em diferentes posições do rosto. 49
Figura 11 – Amplificação de Video Euleriana 50
Figura 12 – Sinais obtidos pelas médias das intensidades de cada canal em cada quadro. 50
Figura 13 – Sinais filtrados. 51
Figura 14 – Resultado obtido pela algoritmo 51
Figura 15 – Resultado obtido pela algoritmo (a esquerda) e estimação apresentada pelo oxímetro de pulso de dedo (a direita) 52
Figura 16 – Resultado obtido pela novo algoritmo proposto 60
Figura 17 – Resultado obtido pela novo algoritmo proposto 61

Lista de tabelas

Tabela 1 – Resultado obtido pelo algoritmo variando a amplificação 53
Tabela 2 – Resultado obtido pelo algoritmo variando a amplificação, considerando a média do erro percentual para a Luz Branca, utilizando o espaço de cor L*a*b*. 53
Tabela 3 – Resultado obtido pelo algoritmo variando a amplificação, considerando o Desvio padrão do erro médio percentual para a Luz Branca, utilizando o espaço de cor $L^*a^*b^*$. 54
Tabela 4 – Média do erro percentual. 54
Tabela 5 – Desvio padrão do erro percentual. 55
Tabela 6 – Média do erro percentual considerando diferentes pontos faciais. 55
Tabela 7 – Desvio padrão do erro percentual considerando diferentes pontos faciais. 55
Tabela 8 – Erro percentual considerando diferentes filtragens. 56
Tabela 9 – Erro percentual considerando diferentes amplificações. 56
Tabela 10 – Erro percentual considerando amplificações aplicadas sobre todos os coeficientes e sobre apenas os coeficientes de aproximação. 57
Tabela 11 – Erro percentual considerando diferentes tipos de wavelets. 57
Tabela 12 – Erro percentual considerando a variação da amplificação para a luz 3, Laranja. 58
Tabela 13 – Erro percentual considerando a aplicação do EVM com as pirâmides Gaussianas e com a transformada de Wavelets. 58
Tabela 14 – Erro médio percentual encontrado ao ser aplicado os quatro tipos de treino. 59
Tabela 15 – Erro médio percentual encontrado ao ser aplicado o treino geral com os trinta e seis vídeos. 60
Tabela 16 – Resultado obtido com a variação da taxa de quadros por segundo . . . 62
Tabela 17 – Resultado do F1 score para os diferentes conjuntos de vídeo. 63
Tabela 18 – Resultado do IoU para os diferentes conjuntos de vídeo. 64

Lista de abreviaturas e siglas

BPM	Batimentos por minuto
ECG	Eletrocardiógrafo
FFT	Transformada Rápida de Fourier
HSV	Espaço de cores onde os níveis de matiz (H), saturação (S) e valor (V) formam as demais cores
KLT	Kanade-Lucas-Tomasi
$L*a*b*$	Espaço de cores onde os níveis lumina (L) e crominância (a* e b*) formam as demais cores
PPG	Fotopletismografia
ICA	Análise de componentes independentes
PH	Potencial Hidrogeniônico
ROI	Região de Interesse
RGB	Espaço de cores onde as cores vermelha (R), verde (G) e azul (B) se combinam para formar as demais cores
SA	Sinoatrial
SpO_2	Saturação de Oxigênio
YIQ	Espaço de cores onde os níveis lumina (Y) e crominância (I e Q) formam as demais cores
$YCbCr$	Espaço de cores onde os níveis lumina (Y) e crominância (Cb e Cr) formam as demais cores
db	Daubechies
fps	*Frames per second*

Lista de símbolos

π	Constante de valor 3.1415926535...
Σ	Somatório
α	Coeficiente angular da equação linear de melhor ajuste
β	Coeficiente linear da equação linear de melhor ajuste
Δ	Diferença entre dois valores
$\mu_{Canal_Vermelho}$	Média das intensidades do canal vermelho
μ_{Canal_Azul}	Média das intensidades do canal azul
σ_{Canal_Azul}	Desvio padrão das intensidades do canal azul
$\sigma_{Canal_Vermelho}$	Desvio padrão das intensidades do canal vermelho
μ_p	Viscosidade
Ω	Resistência
$^\circ$	Graus
BPM_{max}	Maior frequência cardíaca possível
BPM_{min}	Menor frequência cardíaca possível
L	Comprimento
F	Fluxo
r	Raio do tubo
P	Pressão
Hz	Hertz
TR	Taxa R
W	Vetor que armazena as médias do canal verde após a aplicação da janela de Hamming
n	Posição no vetor com as médias originais do canal verde
N	Número de amostras no vetor

BPM	Frequência Cardíaca
sat	Saturação de oxigênio
FPS	Quadros por segundo
R	Canal Vermelho
G	Canal Verde
B	Canal Azul
Y	Canal de luminância
I	Canal de Crominância
Q	Canal de Crominância
V_i	I-ésimo elemento do vetor do sinal vermelho
IV_i	I-ésimo elemento do vetor do sinal infravermelho
$\bar{V}$	Média dos elementos do vetor do sinal vermelho
$I\bar{V}$	Média dos elementos do vetor do sinal infravermelho
$Pico$	Valor da Intensidade no Pico do Sinal
N_{picos}	Número de Picos no sinal
N_{pdm}	Números de Picos do sinal acima da média da intensidades de pico
$P_{Média}$	Valor Médio dos Picos
t	Tempo
f	Frequência
f_{pico}	frequência do maior pico da Transformada de Fourier no intervalo considerado
VP	Verdadeiro positivo
VN	Verdadeiro negativo
FP	Falso positivo
FN	Falso negativo
IoU	Índice de Jaccard

Pr	Precisão
Rc	*Recall*
$F1_{score}$	*F1 score*
B	Conjunto B
A	Conjunto A
$\Delta Pico_i$	Intervalo de tempo entre dois determinados picos
$\Delta Pico_{média}$	Média do tempo entre os picos
Idx_fft_min	Índice da Transformada de Fourier correspondente à menor frequência cardíaca possível
Idx_fft_max	Índice da Transformada de Fourier correspondente à maior frequência cardíaca possível
Idx_{FFT}	Índice da Transformada de Fourier correspondente à frequência do maior pico da FFT(Transformada Rápida de Fourier)
Idx_{pico_i}	Índice do i-ésimo pico
$Idx_{pico_{i-1}}$	Índice do pico que antecede o i-ésimo pico
f_{card}	Frequência cardíaca
i	i-ésimo elemento

Sumário

1 INTRODUÇÃO 23
1.1 Objetivos Gerais 24
1.2 Objetivos Específicos 25

2 ANATOMIA E FISIOLOGIA HUMANA DO SISTEMA CARDÍACO 27
2.1 Sistema Circulatório 27
2.2 Coração 28
2.2.1 Atividade Elétrica 28
2.3 Saturação do Oxigênio 30
2.4 Eletrocardiograma 30
2.5 Pletismografia 33
2.5.1 Fotopletismografia 33

3 ESTIMAÇÃO ÓTICA DA FREQUÊNCIA CARDÍACA E DA SATURAÇÃO DE OXIGÊNIO 35
3.1 Detecção Facial 35
3.2 Amplificação de Vídeo Euleriana 36
3.3 Obtenção da Frequência Cardíaca e da Saturação de Oxigênio . . . 37
3.4 Espaço de Cores 38
3.5 Aprendizado de Máquina 39

4 METODOLOGIA 41
4.1 Etapas Iniciais do Algoritmo 41
4.2 Algoritmo para Testes Finais 43
4.2.1 Programação em Paralelo 43
4.2.2 Aplicação de Métodos Alternativos de Cálculo para Frequência Cardíaca . . 44
4.2.3 Método Alternativo para o Cálculo da Saturação de Oxigênio 45
4.2.4 Espaço de Cor 45
4.2.5 Transformada de Wavelets Bidimensional 45
4.2.6 Aprendizado de Máquina 46
4.3 Materiais 46

5 RESULTADOS E DISCUSSÃO 49
5.1 Algoritmo Inicial 49
5.2 Aplicação de melhorias no Algoritmo 52
5.2.1 Seleção do melhor método de cálculo. 52

5.2.2 Seleção do melhor espaço de cor 52
5.2.3 Seleção do Melhor Conjunto de Pontos Faciais 54
5.2.4 Aplicação da transformada de Wavelets Bidimensional 55
5.2.5 Aplicação do aprendizado de máquina. 58
5.2.6 Testando o efeito do resultado para diferentes taxas de quadro por segundo. 60
5.2.7 Testes de Validação 62

6 CONCLUSÃO 65

REFERÊNCIAS 69

1 INTRODUÇÃO

O coração é, reconhecidamente, um dos órgãos mais importantes do corpo humano. Ao longo dos anos, a interpretação sobre o papel deste órgão foi se transformando. No Egito Antigo, ele era visto como o centro da vida material e espiritual. Entre os Hebreus, era considerado o centro da vida psicológica, moral e espiritual. Atualmente, ele é reconhecido como um órgão indispensável, responsável pelo transporte de oxigênio e nutrientes para os tecidos, além de contribuir para a remoção de resíduos, como o dióxido de carbono.

Apesar de sua função ter sido interpretada de formas diversas ao longo dos séculos, o coração continua a ser objeto de intenso estudo, especialmente na área médica, devido à sua importância.

Essa área estuda diversos problemas de saúde humana, como a doença arterial coronariana, em que o fluxo de sangue para o coração é parcial ou totalmente bloqueado pelo acúmulo de gordura nas artérias; as arritmias, caracterizadas por batimentos cardíacos acelerados (taquicardia) ou lentos (bradicardia); e a síncope vasovagal, em que o indivíduo desmaia devido à queda da pressão arterial e à diminuição dos batimentos cardíacos.

Diversas técnicas podem ser empregadas para a aquisição de sinais cardíacos, sendo o ECG (eletrocardiograma) a mais conhecida. O ECG registra a atividade elétrica do coração por meio de eletrodos posicionados em diferentes partes do corpo. As variações de tensão, geradas pelas células cardíacas excitáveis, produzem uma série de ondas que representam os batimentos cardíacos. A partir do intervalo entre dois picos R da onda, é possível calcular a frequência cardíaca instantânea. Esses dados permitem o diagnóstico de doenças associadas a distúrbios na atividade elétrica cardíaca (SÖRNMO; LAGUNA, 2005).

Embora não-invasivo, o ECG pode ser bastante incômodo em determinadas situações, causando desconforto ao paciente. No caso de cardiopatas com marcapasso, por exemplo, a necessidade de realizar múltiplos exames, incluindo variações do eletrocardiograma, pode ser desconfortável, seja por alergias a algum material, rejeição ao exame ou insatisfação com a repetição anual do procedimento. Esse cenário pode levar o paciente a evitar novos exames, comprometendo o acompanhamento médico e aumentando os riscos à saúde. Como muitos dos pacientes diagnosticados com arritmias e outras doenças cardíacas são de idade avançada, há uma demanda crescente por procedimentos mais confortáveis, precisos e que não afastem os pacientes. Além disso, em situações de pandemia, muitos médicos preferem manter esses pacientes longe de ambientes hospitalares,

pois fazem parte de grupos de risco, sendo mais vulneráveis.

Neste contexto, uma alternativa não-invasiva, menos incômoda e que possa ser aplicada de forma remota torna-se almejável, tanto para profissionais da saúde quanto para paciente. Uma alternativa que vai de encontro a estes pontos é a fotopletismografia que, baseada no princípio pletismográfico e sob o auxílio de luzes infravermelhas e conjuntos de fotorreceptores, calcula a variação do volume de sangue (TUSMAN; BOHM; SUAREZ-SIPMANN, 2017). Esta variação provoca mudanças sutis na tonalidade da pele, imperceptíveis ao olho humano e que podem ser usadas como parâmetro para se estimar a frequência cardíaca (WU et al., 2012a). Seguindo esta lógica, alguns estudos vêm sendo desenvolvidos buscando obter a frequência cardíaca de pacientes através de segmentos de vídeos do mesmo (LIMA; GUSMÃO; COSTA, 2019).

Fátima e Cesar (ROSA; BETINI, 2020) realizaram um estudo onde, com um auxílio de uma câmera Raspberry Pi e um microprocessador Raspberry Pi 3, era possível aferir a frequência cardíaca e a saturação de oxigênio no sangue . O sistema realizava uma calibração prévia com cada voluntário e com os dados coletados estimava as informações desejadas comparando estas informações com as obtidas a partir de um oxímetro de referência. Os resultados foram promissores, principalmente do ponto de vista da saturação de oxigênio no sangue, no entanto, apesar de aceitável, a aferição da frequência cardíaca precisava ser aprimorada.

Cobos-Torres e Abderrahim também realizaram estudo para a obtenção da SpO_2, realizando em tempo real e realizando 100% do processamento online (COBOS-TORRES; ABDERRAHIM, 2017). O trabalho de Poh e Poh também realizava o monitoramento em tempo real para obtenção da frequência cardíaca, com auxílio do software proprietário "Cardiio Heart Rate app" (POH; POH, 2017). De forma semelhante, Nowara, McDuff e Veeraraghavan (NOWARA; MCDUFF; VEERARAGHAVAN, 2021) analisaram a interferência da compressão de vídeo na estimação do pulso a partir de vídeos. Outros estudos serão mencionados de forma mais detalhada ao longo deste estudo.

Dada a necessidade de adquirir os sinais cardíacos de forma remota e menos incômoda, sem perder a coerência, a validade e sem aumentar muito os custos, o presente trabalho busca aferir a frequência cardíaca com base em vídeos das faces dos pacientes.

1.1 Objetivos Gerais

Implementar um algoritmo capaz de estimar a frequência cardíaca, bem como a saturação de oxigênio, a partir de dados extraídos de segmentos de vídeo, utilizando conhecimentos de fotopletismografia e processamento de imagem.

1.2 Objetivos Específicos

1. Realização da detecção facial para cada quadra da sequência de vídeo de entrada
2. Processamento dos quadros de segmentos de vídeos para realçar as mudanças imperceptíveis a olho nu
3. Filtragem do sinal para eliminação de ruídos
4. Aplicação de cálculos eficiente para estimar os parâmetros desejados
5. Otimização dos resultados encontrados por meio de algorítimos de aprendizado de máquina
6. Obter medidas frequências cardíacas e saturação de oxigênio com um erro médio percentual baixo

2 ANATOMIA E FISIOLOGIA HUMANA DO SISTEMA CARDÍACO

2.1 Sistema Circulatório

Responsável pelo transporte de nutrientes, pela defesa do organismo, regulação da temperatura, estabilização do PH e pela homeostase, o sistema circulatório é um sistema fechado composto por vasos sanguíneos e pelo coração que bombeia o sangue para o corpo. Este sistema pode, também, ser dividido em três partes: sistema sanguíneo (composto por veias, artérias, arteríolas, vênulas, capilares e o coração) , sistema linfático (composto por capilares linfáticos, vasos linfáticos, troncos linfáticos e órgãos linfoides) e órgãos hematopoiéticos (medula óssea, baço e timo).

Segundo Nivea Cristina Santos, o coração, órgão central desse sistema, funciona como uma bomba contrátil. Esta estrutura é composta pelo tecido muscular estriado cardíaco, o qual é formado pelo endocárdio, miocárdio e epicárdio (SANTOS, 2018). Sua morfologia interna está ligada às quatro câmaras que têm sua origem na cavidade cardíaca e pode ser visualizada na Figura 1.

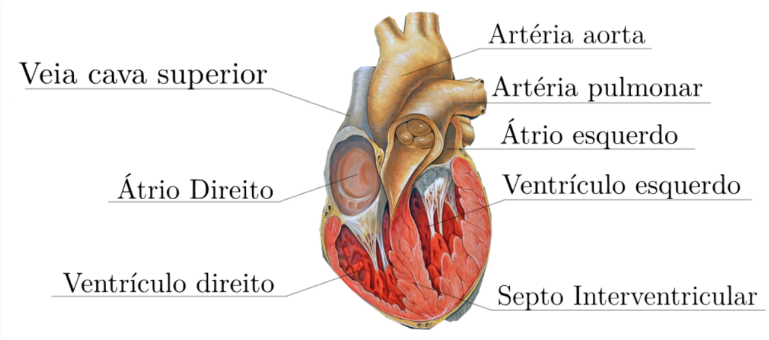


Figura 1 – Morfologia interna do coração

Fonte: Adaptado de (LYNCH, 2006)

Através da veia cava superior e inferior e da veia coronária, o sangue venoso de todo organismo é transportado a uma destas câmaras, também chamadas de septos, conhecida como átrio direito. Através do ventrículo direito, que possui uma musculatura mais espessa, o sangue venoso é transportado para os pulmões através da artéria pulmonar, que se divide em artéria pulmonar direita e esquerda cada uma ligada a um pulmão.

Do lado esquerdo do coração, as veias pulmonares desembocam no átrio esquerdo, transportando o sangue arterial proveniente dos pulmões. Esse sangue é então levado ao ventrículo esquerdo, a cavidade com a musculatura cardíaca mais espessa, de onde sai a artéria aorta.

Também fazem parte da morfologia interna do coração a valva tricúspide, valva bicúspide, valva pulmonar e valva aórtica. A primeira, localizada entre o átrio direito e o ventrículo direito, permite a passagem de sangue do átrio para o ventrículo e impede o refluxo de sangue no sentido inverso. A segunda, localizada entre o átrio e ventrículo esquerdo, funciona de maneira análoga à primeira, porém do lado esquerdo do coração. A Valva pulmonar, localizada entre o ventrículo direito e a artéria pulmonar, permite a passagem de sangue no sentido do ventrículo para as artérias e impede o refluxo no sentido contrário. A valva aórtica funciona de forma análoga a valva pulmonar e está situada entre o ventrículo esquerdo e a artéria aorta.

Desta forma, com auxílio das veias e artérias, o sangue é transportado por todo o organismo transportando os nutrientes necessários. Para o estudo em questão, o foco será mantido no coração e nos seus efeitos no sistema circulatório.

2.2 Coração

2.2.1 Atividade Elétrica

Conforme Sörno e Laguna, o miocárdio, citado anteriormente, é a parede do coração composta por células musculares. Estas realizam uma força mecânica, ativa por um impulso elétrico que percorre todo o coração graças a uma rede de células especializadas, durante a contração (SÖRNMO; LAGUNA, 2005). O ciclo cardíaco é gerado, então, quando um impulso elétrico percorre esse sistema de condução iniciando uma série de eventos mecânicos que começam no átrio direito, onde o sangue chega vindo de todas as veias, exceto as dos pulmões. Essa contração conduz o sangue para o ventrículo direito, o qual, após ser preenchido, se contrai e transporta o sangue para os pulmões. Nos pulmões, o excesso de dióxido de carbono é substituído por oxigênio e, em seguida, o sangue retorna para o coração pelo átrio esquerdo. O átrio esquerdo se contrai e o sangue é direcionado ao ventrículo esquerdo, de onde será bombeado para todo o corpo através dos vasos arteriais e dos capilares. Após percorrer o corpo, o sangue arterial, agora rico em dióxido de carbono, retorna ao coração, dando início a um novo ciclo.

O ciclo cardíaco é composto por um momento de ativação e de recuperação, correspondendo a despolarização e repolarização, do ponto de vista elétrico, e a contração e relaxamento, do ponto de vista mecânico. Iniciando no nó sinoatrial, grupo de células marcapasso, um impulso elétrico é conduzido através do miocárdio por meio da des-

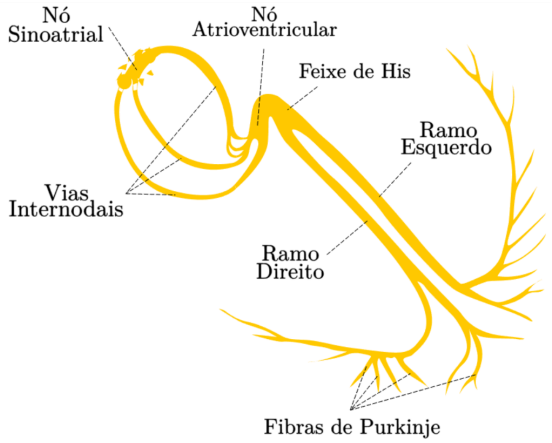


Figura 2 – Sistema de Condução Elétrica do Coração

Fonte: Adaptado de (MADHERO88, 2009)

polarização, que se dá graças a rápida variação do potencial da célula que estimula a despolarização das células vizinhas e, assim, conduz o impulso. A despolarização é rapidamente seguida da repolarização e esse processo permite que a contração e relaxamento ocorram de forma coordenada no tempo correto. Esse intervalo entre a contração do átrio e a contração do ventrículo ocorre pelo fato de que o impulso que se inicia no átrio passa pelo nó atrioventricular, onde é atrasado graças a condução mais lenta existente nesta área do tecido muscular. O impulso atravessa a parede entre os dois ventrículos através do feixe de His que conecta eletricamente os átrios e o nó Atrioventricular. Esse feixe se subdivide em outros dois feixes de condução rápida, um no ventrículo esquerdo e outro no ventrículo direito, se ramificando ainda nas chamadas fibras de Purkinje. A propagação nesta região deve ser rápida em virtude da largura dos ventrículos, para que a contração seja unificada.

O nó sinoatrial funciona portanto como um marcapasso natural que determina a frequência cardíaca, possuindo a maior frequência de marcapasso em relação às demais células que seguem em sincronia. Os batimentos cardíacos no entanto, não são determinados apenas pela taxa de descarga do nó SA, mas também pelas informações do exterior que são mediadas pelo sistema nervoso autônomo, onde um aumento nas atividades simpáticas, responsável pelas mudanças no organismo em situações de alerta, estresse ou emergência, provocam um aumento na frequência cardíaca, enquanto um aumento nas atividades parassimpáticas, responsável por retornar o estado de calma em que se encontrava antes da situação de alerta, provoca uma redução na mesma.

O limite superior dos batimentos cardíacos está ligado à taxa de descarga pelo tempo em que um célula está eletricamente inativa, sendo este limite uma taxa teórica máxima de 240-300 batimentos por minuto. No entanto, na prática, raramente a frequência no nó sinoatrial passa de 220 batimentos por minuto e esta frequência baixa à medida

que o indivíduo envelhece.

2.3 Saturação do Oxigênio

Ainda sobre o sistema circulatório, é necessário abordar a saturação de oxigênio. Esta indica quantidade de oxigênio que circula no sangue, isto é, a relação entre a quantidade de hemoglobinas ligadas a moléculas de oxigênio e a quantidade de hemoglobinas sem essa ligação. Esta relação normalmente se encontra numa faixa acima de 95%. Contudo algumas doenças respiratórias e cardíacas são capazes de afetar a saturação, fazendo com que ela caia para até 88%.

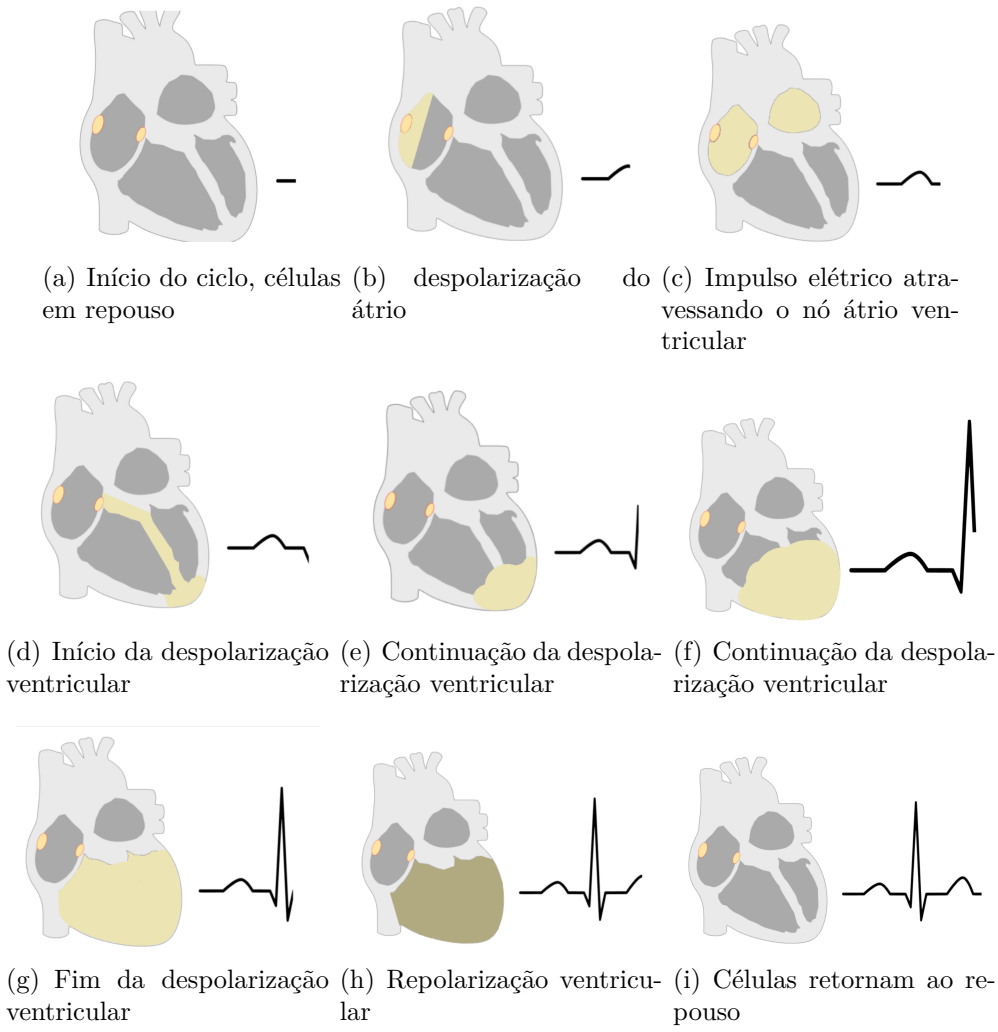
Esta informação, para medicina, auxilia na identificação de doenças e complicações citadas anteriormente, bem como o tratamento destas. Ela também é utilizada para identificar as condições de indivíduos internados. Por ser de extrema importância o transporte de oxigênio, uma queda intensa da saturação pode resultar em alguns sintomas como : falta de ar, tontura, sonolência. No entanto, existe o caso de hipóxia silenciosa, que recentemente, com a pandemia, teve seus casos aumentados em virtude da covid-19. Esta consiste numa queda repentina da saturação, porém esta não apresenta nenhum sintoma, mesmo que possa chegar a taxas de 60%.

Para a análise destes dados dois exames são utilizados com mais frequência, sendo eles a gasometria, em que uma amostra de sangue é coletada para se aferir a concentração de oxigênio e outras informações, e o oxímetro, que mede o nível de oxigênio no sistema circulatório que passa pelo dedo através de um sensor de luz. Em geral o primeiro exame é utilizado para se obter informações mais detalhadas e precisas, porém isso não afeta a veracidade dos dados obtidos pelo oxímetro. O presente estudo irá considerar também a saturação se baseando em resultados obtidos por oxímetros. Porém, para que o entendimento não seja prejudicado, mais a frente será realizado uma revisão acerca da Pletismografia.

2.4 Eletrocardiograma

Sörno e Laguna também explicam que as ondas dos sinais ECG podem ser compreendidas levando-se em consideração que, dentre as milhares de células do coração, apenas um grupo despolariza simultaneamente, sendo representados por um vetor associado a uma fonte de corrente dipolo, descrevendo a variação no tempo da magnitude, orientação e posição do dipolo. Estes vetores podem ser somados, resultando num vetor dominante que descreve a direção principal do impulso elétrico (SÖRNMO; LAGUNA, 2005).

No início do ciclo, todas as células cardíacas se encontram em repouso, o que resulta numa linha horizontal na onda do sinal ECG, como se pode ver na Figura 3(a). Durante



(a) Início do ciclo, células em repouso

(b) despolarização do átrio

(c) Impulso elétrico atravessando o nó átrio ventricular

(d) Início da despolarização ventricular

(e) Continuação da despolarização ventricular

(f) Continuação da despolarização ventricular

(g) Fim da despolarização ventricular

(h) Repolarização ventricular

(i) Células retornam ao repouso

Figura 3 – Sinal ECG registrado posicionado ao lado do coração.
Fonte: Adaptado de (YITZHAKNAT, 2022)

a despolarização dos átrios o vetor dominante é direcionado para baixo em direção ao nó atrioventricular, resultando numa onda registrada com polarização positiva (Figura 3(b)). Quando se finaliza a despolarização do átrio, o sinal ECG volta para um valor constante (isoelétrica), Figura 3(c). A despolarização do nó atrioventricular e do feixe de His inicia assim que a do átrio termina, porém, não é detectado pelo ECG por causa da condução lenta nesta região. Por outro lado, o sinal registrado na despolarização do ventrículo é muito maior que a atrial, por apresentar maior massa muscular, Figura 3(d). A onda apresenta polarização negativa pois se inicia na parede entre o ventrículos de forma que o vetor determinante se direciona para longe do eletrodo. Devido a rápida condução na região, a onda terá curta duração e será estreita.

Ao longo da despolarização do ventrículo o vetor dominante vai gradativamente mudando sua direção para o eletrodo, porém, ao final da despolarização, o dominante se encontra apontando para longe do eletrodo. Essa mudança mais lenta se dá pelo fato de que as paredes do ventrículo esquerdo são três vezes mais espessas que a do direito e,

portanto, demora mais para despolarizar, Figura 3(e) - 3(g). Paralelo à despolarização ventricular, ocorre a repolarização atrial, porém este sinal é formado em sua maior parte pela despolarização, pois este apresenta amplitudes muito maiores. Após a despolarização ventricular, o sinal ECG retorna a curva isoelétrica até que se inicie a repolarização dos ventrículos, onde o vetor dominante se comporta de forma semelhante ao ocorrido durante a despolarização ventricular, dando origem uma polarização positiva registrada pelo eletrocardiograma, Figura 3(h). Após a despolarização, as células cardíacas retornam ao repouso(SÖRNMO; LAGUNA, 2005).

Para completar a análise de ECG, vale ressaltar a nomenclatura convencional de cada onda. Chama-se onda P a onda correspondente à despolarização atrial. O complexo QRS é correspondente à despolarização dos ventrículos, onde a onda Q representa a deflexão negativa, a onda R a deflexão positiva e a S a negativa que a sucede. Neste ponto, acrescenta-se um breve hiato para o destaque de algumas informações pertinentes. O intervalo RR, isto é, o intervalo entre uma onda R e a seguinte, corresponde ao tempo do ciclo ventricular e é fundamental para a determinação de ritmos cardíacos, sendo utilizado para o estudo de arritmias e variação da frequência cardíaca. Vale relembrar que, paralelo à formação da onda QRS, também ocorre a repolarização dos átrios, porém este sinal é mascarado pela despolarização dos ventrículos. Por fim, a onda T indica a repolarização ventricular.

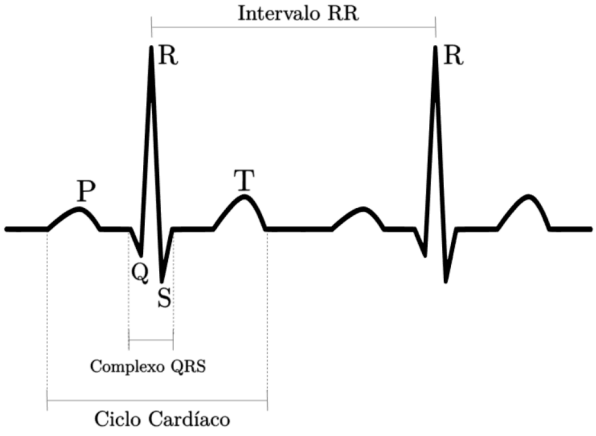


Figura 4 – Definições das ondas do ciclo cardíaco e suas durações.
Fonte: Adaptado de (YITZHAKNAT, 2022)

Para que estes sinais sejam captados de forma eficiente, são necessários múltiplos eletrodos que serão dispostos sobre a superfície do corpo, com contato, além de inúmeras derivações, considerações e tratamento das informações. O que torna o exame de difícil e incômoda realização, quando executado por longas durações.

2.5 Pletismografia

A pletismografia é um método mais sofisticado que visa avaliar o funcionamento pulmonar, baseando-se em informações acerca da pressão, da resistência pulmonar, da resistência nas vias aéreas e na resistência respiratória total, de modo que estas resistências são medidas para condições de fluxo laminar e estruturas rígidas, considerando o modelo baseada na Lei de Poiseuille:

$$P = \mu_p \frac{LF}{\pi r^4} \tag{2.1}$$

Nesta equação a pressão é associada com a razão entre fluxo aéreo e a área da seção transversal da via aérea. Com isso a Resistência pode ser calculada com base na aplicação da Lei de Ohm, onde:

$$\Omega = \frac{P}{F} \tag{2.2}$$

A resistência das vias aéreas seria então definida como a pressão necessária para produzir um fluxo de $1L/s$ nos pulmões. Assim como nos estudos de circuitos elétricos, a condutância das vias aéreas seria o inverso da resistência e corresponderia ao fluxo gerado por unidade de pressão (ALBERTO; PEREIRA; MOREIRA, 2002).

Assim, considerando que a pressão medida é a mesma força propulsora que causou o fluxo antes da oclusão e que, na ausência de fluxo, a pressão em um ponto dentro de um tubo rígido preenchido por um gás, ou fluido, será distribuída uniformemente através do tubo, no exame pletismográfico a pressão alveolar será estimada após a oclusão de um obturador situado ao nível da boca em um ambiente climatizado e totalmente fechado. Este, apesar de desenvolvido em nível tecnológico e capaz de obter informações acerca de doenças, como a enfisema, asma e insuficiência cardíaca pulmonar, ainda pode ser considerado incômodo e invasivo.

2.5.1 Fotopletismografia

Assim como o eletrocardiograma, a fotopletismografia é um método não-invasivo de aquisição de sinais biológicos que, baseado em princípios da pletismografia, é capaz de estimar a saturação de oxigênio e a frequência cardíaca a partir do volume de sangue.

Assim como no ECG, através do formato de onda é possível obter informações acerca do paciente. No entanto, o formato de onda é diferente, e as análises das ondas P, QRS e T não são válidas aqui. A nova onda em questão tem suas informações contidas na altura, largura e no nó dicrótico (TUSMAN; BOHM; SUAREZ-SIPMANN, 2017), que sofrem alterações em seus formatos conforme a presença de doenças cardíacas.

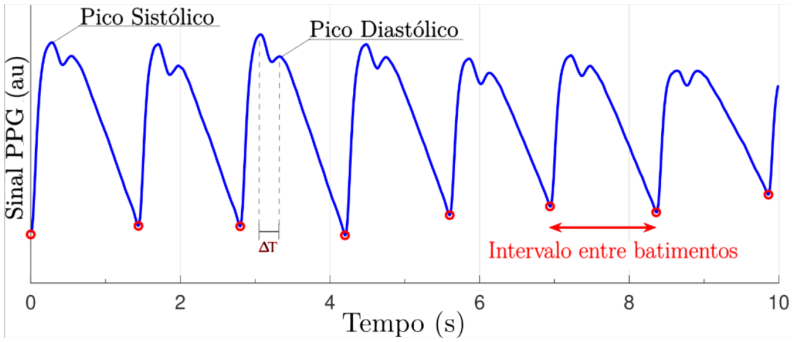


Figura 5 – Onda da fotopletismografia

Fonte: Adaptado de (PETERHCHARLTON, 2020)

A captação dos sinais se dá através de sensores óticos que podem ser posicionados de duas formas distintas. Os sensores podem ser posicionados em lados opostos, de forma que um seja um emissor com luzes infravermelha e vermelha e o outro um receptor no lado oposto. A oxiemoglobina é capaz de absorver as luzes dos sensores emitidos, portanto, quanto menor a luz incidente no receptor, maior a saturação. Por fim os sensores ainda podem ser colocados do mesmo lado, sendo geralmente utilizados, quando presilhas são menos acessíveis.

Na onda PPG(*photoplethysmogram*) as variações em altas frequência são provocadas pelas mudanças no fluxo sanguíneo arterial, enquanto as de baixa frequência são causadas pelas variações no sangue venoso e de outros componentes do tecido.(MEJIA-MEJIA et al., 2021)

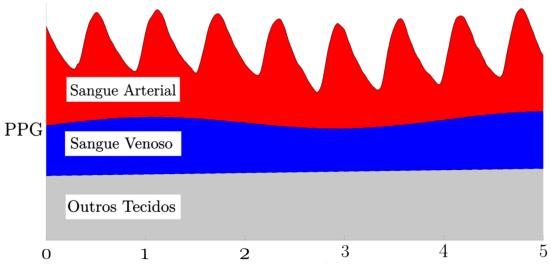


Figura 6 – Ondas do sinal PPG.

Fonte: Adaptado de (PETERHCHARLTON, 2021)

3 ESTIMAÇÃO ÓTICA DA FREQUÊNCIA CARDÍACA E DA SATURAÇÃO DE OXIGÊNIO

Com os conceitos sobre o sistema circulatório cardíaco resumidos juntamente com a breve explicação sobre os meios mais usuais para se aferir tanto a frequência cardíaca, como a saturação de oxigênio do corpo, serão, a seguir, apresentadas as ferramentas utilizadas no presente trabalho. Neste capítulo também será apresentado o estado da arte do tema, que exemplifica e reforça a eficácia do método. O método apresentado busca, assim como os anteriores, identificar a frequência cardíaca do usuário, além de buscar aferir também a saturação de oxigênio, no entanto, não mais por objetos que estão em contato com a pele do indivíduo e sim por segmentos de vídeo.

3.1 Detecção Facial

Dentro do contexto apresentado, um novo método se propõe a medir a frequência cardíaca e a saturação de oxigênio através de segmentos de vídeo. Apesar do objetivo final ser o mesmo, diversas são as técnicas que podem ser aproveitadas, no entanto, a grande maioria, necessita de uma região de interesse (ROI). Esta é selecionada para isolar as características visuais da face do indivíduo analisado, que são utilizadas para medir a frequência cardíaca e a saturação de oxigênio, podendo variar desde uma região da testa, regiões da bochecha, pontos do rosto ou ainda uma combinação entre esses. No estudo do Arnoldo *et al.* (LIMA; GUSMÃO; COSTA, 2019), por exemplo, o algoritmo Viola Jones (VIOLA; JONES, 2001) é utilizado para que o rosto seja identificado no primeiro quadro do vídeo e assim, a região de interesse pode ser definida. No entanto, o autor identificou que este algoritmo apresentava um alto custo computacional, além de só conseguir identificar o rosto humano se este estivesse na vertical, sem inclinação, de frente para a câmera, com vistas completas, sem nenhuma obstrução e bem iluminadas. Dito isso, para o rastreamento, o algoritmo Kanade-Lucas-Tomasi (KLT) foi utilizado (SHI; TOMASI, 1994). O KLT aumenta a velocidade de rastreio, pois utiliza os pontos característicos do primeiro quadro, onde houve a detecção facial pelo Viola-Jones, e a partir destes pontos, detecta a face no quadro seguinte, reduzindo o tempo e o custo computacional. Estas não são as únicas formas de se realizar a detecção facial. Outros métodos, como a utilização de redes neurais, são viáveis e por vezes mais eficientes. Com o rosto identificado, é importante definir a região de interesse, pois pontos como os olhos,

a boca, os pelos faciais e o cabelo podem acrescentar ruídos indesejados ao sinal a ser extraído. Uma forma eficaz de se definir a região de interesse é aproveitando as *landmarks* da detecção facial de bibliotecas como a *Mediapipe* (LLC, 2022). Basicamente, ao realizar a detecção do rosto, alguns pontos são definidos e rastreados. Dessa forma, ainda que o usuário incline o rosto ou vire-o, dentro de uma tolerância, o rosto ainda será detectado e rastreado.



Figura 7 – Detecção Facial com *Landmarks* Selecionados Destacados

Assim é possível usar estes pontos para criar a região de interesse ou até mesmo utilizá-los como pontos de interesse e, assim remover informações dos canais de interesse.

3.2 Amplificação de Vídeo Euleriana

Com o rosto identificado e com as regiões de interesse determinadas, os próximos passos a serem realizados são o tratamento da imagem e a filtragem, para que a variação da coloração facial seja realçada para se estimar os dados desejados. Uma forma de fazer isto é utilizando a Magnificação Euleriana de Vídeo (WU et al., 2012a) . Para este algoritmo, primeiro deve-se construir a pirâmide Gaussiana com cada quadro do vídeo, isto é, deve-se suavizar cada imagem, por meio do filtro Gaussiano e, em seguida, sub amostrar a imagem para cada nível construído. O número de níveis usualmente utilizado é 3 ou 4. Todo o tratamento da imagem é realizado nesse último nível da pirâmide.

Primeiro realiza-se a filtragem temporal com filtro ideal na frequência mínima e máxima, sendo esta a faixa de batimentos comum em um ser humano saudável. Com isso

são selecionadas apenas as mudanças que correspondem às alterações na coloração da face do indivíduo promovidos pelo fluxo sanguíneo. Escolhe-se então dois fatores, Amplificação e Atenuação, para realçar as tais mudanças. Com tais configurações, reconstrói-se a pirâmide Gaussiana através da sobre amostragem das imagens modificadas. O quadro do último nível possui a mesma resolução que a imagem original do vídeo, porém filtrada e amplificada. O último passo é somar à imagem original a imagem filtrada, obtendo-se a amplificação da alteração da coloração facial à medida que o coração bombeia o sangue para o corpo. Nesta etapa, tanto os filtros quanto as amplificações e pirâmides podem ser substituídas por outras além das mencionadas, em busca de um resultado melhor, baseado no vídeo em que se está utilizando.

3.3 Obtenção da Frequência Cardíaca e da Saturação de Oxigênio

Para a última etapa da proposta, resta utilizar os dados obtidos após a filtragem e amplificação. Os métodos utilizados para a realização do cálculo podem variar, conforme o desafio é abordado. Rosa e Betini utilizaram duas abordagens diferentes para cada parâmetro em seu estudo (ROSA; BETINI, 2020). Em caso gerais, quando a imagem está no espaço de cor RGB, as informações acerca da frequência cardíaca se encontrarão no canal verde (G), enquanto as informações da saturação de oxigênio estarão presentes nos canais vermelho e azul (R e B, respectivamente), que equivalem a luz vermelha e infravermelha, respectivamente, segundo alguns estudos, (SCULLY et al., 2012) , (TARASSENKO et al., 2014), (BAL, 2015). No artigo mencionado, de Alessandra *et al.*, os autores criam três vetores que correspondem às médias de cada canal (R, G e B), além de outros dois que correspondem aos desvios padrão dos canais R e B. Para o cálculo da frequência cardíaca, primeiro realiza-se uma filtragem com uma janela *hamming*, para evitar ruídos de alta frequência, multiplicando cada valor do vetor com as médias pela equação:

$$W[n] = 0,54 - 0,46 \cdot cos(\frac{2 \cdot \pi \cdot n}{N-1}) \tag{3.1}$$

Em seguida são aplicadas as transformações de Fourier e é calculado o espectro de potência. Por fim, aplica-se um filtro ideal e calcula-se o índice máximo e mínimo (Equações3.3 e 3.2), onde BPM_{min} e BPM_{max} correspondem a menor e maior frequência cardíaca possível, Idx_fft_min e Idx_fft_max são os índices da Transformada de Fourier correspondentes à menor e a maior frequência cardíaca possível, respectivamente, FPS corresponde a taxa de quadros por segundo e N o número de amostras no vetor.

$$Idx_fft_min = \frac{BPM_{min} \cdot N}{FPS \cdot 60} \tag{3.2}$$

$$Idx_fft_max = \frac{BPM_{max} \cdot N}{FPS \cdot 60} \tag{3.3}$$

Com estas frequências, detecta-se a frequência cardíaca como sendo o índice do maior pico da FFT, dentro dos limites determinados. Assim a frequência cardíaca pode ser estimada através da equação

$$f_{card} = \frac{Idx_{FFT} \cdot FPS \cdot 60}{N} \tag{3.4}$$

Por fim para a saturação de oxigênio é utilizado o desvio padrão da intensidade dos canais Azul e Vermelho em cada quadro, além do valor médio. Assim obtém-se a taxa R, relação normalizada de intensidade de luz vermelha e infravermelha transmitida/absorvida, pela equação 3.5.

$$TR = \frac{\mu_{Canal_Vermelho} \cdot \sigma_{Canal_Azul}}{\mu_{Canal_Azul} \cdot \sigma_{Canal_Vermelho}} \tag{3.5}$$

Com a taxa R, a saturação de oxigênio é obtida pela equação linear 3.6, que utiliza os parâmetros α e β, que são adquiridos por treinamento do dispositivo, para ajustar o cálculo da saturação de oxigênio.

$$sat = \alpha \cdot R + \beta \tag{3.6}$$

3.4 Espaço de Cores

Além das técnicas mencionadas, a seleção do espaço de cor também irá afetar a forma em que se poderá modificar cada quadro em busca das informações necessárias. Usualmente utiliza-se o espaço de cor RGB, onde a imagem é formada a partir de tons de vermelho, verde e azul. Outros canais que podem ser utilizados são os canais **YIQ** (onde **Y** corresponde ao luma e **I** e **Q** a crominância da imagem) e **HSV** (onde **H** é o matiz, a própria cor, **S** a saturação da cor e **V** é o brilho da cor). Dependendo do canal que se escolha, é possível realizar um processamento melhor da imagem ou obter informações mais precisas, como foi realizado anteriormente com o canal **RGB** para obter a saturação de oxigênio pelos canais **R** e **B** e a frequência cardíaca, pelo canal **G** ((ROSA; BETINI, 2020)). Em termos de programação, as imagens ou quadros dos vídeos são, normalmente, representados em canais RGB. Para isso, se faz necessária a mudança do espaço de cor, caso se deseje utilizar o YIQ, HSV ou qualquer outro.

Assim como o YIQ e o HSV, outros espaços de cor que separam informações de luminância e de crominância são os espaços **L*a*b** e **YCbCr**. No caso do YCbCr, as informações de luminância encontram-se presentes no canal **Y**, enquanto as informações de cor se fazem presentes nos canais **Cb** e **Cr**, que se referem às diferenças de cor azul e vermelha, respectivamente, e um valor de referência. Estes canais são muito utilizados para

representação de cores em vídeos digitais. Para o caso do espaço de cor L*a*b, também ocorrerá o desacoplamento da luminância, neste caso, representado pela clareza(**L***) e cor, representado pelo **a***, que corresponde ao vermelho menos verde, e pelo **b***, que corresponde ao amarelo menos azul, conforme explica Gonzalez (GONZALEZ; WOODS, 2009). Estes dois últimos espaços de cor são mais recentes que os mencionados anteriormente, apresentando uma eficácia significativa para a manipulação de imagens.

3.5 Aprendizado de Máquina

Um recurso muito útil, e em grande desenvolvimento atualmente, se trata do uso da inteligência artificial em programas para a resolução de problemas. Diversas são as aplicações atuais em que este recurso tem sido utilizado, quer seja para a automação de automóveis, ou ainda para diagnósticos automáticos(LUDERMIR, 2021).

A ideia por trás da inteligência artificial é resolver problemas empiricamente, com base em dados e informações que permitam o treinamento do programa. Dentro deste assunto, uma das áreas existentes é o aprendizado de máquina, cujo objetivo é fazer com que os softwares melhorem seu desempenho partir de exemplos. Para isso, é necessário um grande volume de dados fornecidos ao algoritmo para que este realize um aprendizado automático, gerando hipóteses a partir destas informações. Quanto maior a qualidade dos dados, melhor será o resultado.

Dentro do aprendizado de máquina, três principais tipos se destacam: aprendizado supervisionado, não-supervisionado e por reforço. O primeiro consiste no treinamento realizado treinamento a partir de dados já classificados, para que o algoritmo consiga realizar as próximas classificações dos exemplos não-rotulados de forma adequada.No segundo tipo, os exemplos não são rotulados, e o próprio algoritmo realiza as classificações, com base em semelhanças identificadas entre os atributos dos dados. No ultimo tipo, o programa não possui informações sobre a resposta certa, mas gera hipóteses e determina se estas foram adequadas ou não, recebendo um estímulo de resposta.(LUDERMIR, 2021)

Considerando o treinamento supervisionado, diversos são os algoritmos existentes para o aprendizado de máquina, sendo alguns exemplos: o aprendizado por regressão linear, por regressão polinomial e por árvore de decisão. O primeiro busca criar um modelo matemático linear que relaciona as variáveis de entrada com as variáveis de saída. No segundo procura criar um modelo polinomial de um determinado grau que relaciona esses pares de entrada e saída. Por outro lado, no terceiro divide os conjuntos de dados em grupos cada vez menores, formando uma estrutura de árvore com nós de decisão, que se ramificam em duas ou mais direções. Os nós folha trazem o resultado final. Ou seja, essa última realiza seu aprendizado e tomadas de decisão por meio de etapas bem definidas.

No contexto em questão, o uso do aprendizado supervisionado é útil e eficaz para

a obtenção de melhores resultados.

4 METODOLOGIA

Nos capítulos anteriores, foram explicados tanto os conceitos físicos ligados ao sistema circulatório quanto sobre os batimentos cardíacos. Estes conceitos ajudam a entender melhor o porquê é possível se aferir a frequência cardíaca e a saturação de oxigênio através de imagens de vídeo, uma vez que os batimentos promovem a circulação do sangue no corpo e esta promove uma alteração sutil na coloração do rosto, alteração esta que é imperceptível a olho nu. Também foram apresentados outros métodos de se aferir tanto a frequência cardíaca quanto a saturação, visando apresentar um comparativo entre os princípios envolvidos nas medições, além de evidenciar o quão tais métodos eram mais invasivos ou desconfortáveis do que o proposto no presente texto. Por fim, para um entendimento melhor dos estudos sobre o assunto, foram apresentados alguns algoritmos utilizados para o mesmo fim, com tratamentos específicos para cada etapa, a fim de fundamentar a escolha realizada durante o desenvolvimento do trabalho. Para isso, propositalmente, foram evidenciados estudos que se aproximassem das decisões tomadas e expressas nas linhas que se seguem. Portanto, neste capítulo será apresentado o trabalho desenvolvido em suas etapas.

4.1 Etapas Iniciais do Algoritmo

A primeira etapa realizada do algoritmo foi a detecção facial, através da biblioteca *mediapipe* (LLC, 2022), onde a face é detectada em cada quadro e nela são selecionadas as *landmarks* desejadas, conforme apresentado na Figura 7. Os índices destas marcações no rosto são armazenadas em uma variável que irá ser utilizada, posteriormente, para obter os canais vermelho, verde e azul.

Com os índices obtidos, inicia-se então o processo de amplificação Euleriana. Antes da construção da pirâmide Gaussiana, realiza-se a conversão do espaços de cor RGB para o espaço YIQ. No algoritmo, fez-se uso da biblioteca *Skimage* (WALT et al., 2014) para a mudança do espaço de cor, porém, é possível realizar a mesma mudança de espaço pelas equações 4.1,4.2 e 4.3.

$$Y = 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B \tag{4.1}$$

$$I = 0,596 \cdot R - 0,275 \cdot G - 0,321 \cdot B \tag{4.2}$$

$$Q = 0,212 \cdot R - 0,523 \cdot G + 0,311 \cdot B \tag{4.3}$$

Com a mudança de espaço, inicia-se então a construção da pirâmide Gaussiana a partir de cada quadro original do vídeo, conforme mencionado anteriormente. No projeto em questão foram utilizados três níveis para a pirâmide Gaussiana. Em seguida aplica-se o filtro passa-faixa temporal no último nível da pirâmide, isto é, aquele mais sub-amostrado. Para esta filtragem optou-se pela faixa de frequências entre $0,8333Hz$ e $1,5$. Normalmente os batimentos cardíacos podem encontrar-se nessa faixa em um adulto normal. Ainda que, neste caso, o indivíduo analisado possua arritmia, ele é portador de marcapasso, o que faz com que seus batimentos sejam regulados pelo aparelho. Após a filtragem, os três canais, de cada imagem do vídeo, são amplificados por um fator de amplificação. Posteriormente, os canais I e Q, ou G e B, caso seja utilizado o espaço RGB, são atenuados por um fator de atenuação. O algoritmo inicialmente utiliza o espaço YIQ, posteriormente sendo alterado, com um fator de atenuação com valor igual a 1, enquanto o fator de amplificação assumia valor igual 30.

Com a amplificação realizada, inicia-se a reconstrução da pirâmide Gaussiana, como o mesmo número de níveis utilizado anteriormente, através da sobre-amostragem, retornando assim a resolução original. Por fim, é realizado novamente, em cada quadro, a mudança do espaço de cor para o espaço RGB. Por fim, somava-se o vídeo amplificado e filtrado ao vídeo original, realçando as mudanças sutis na coloração da face do indivíduo.

Com a amplificação Euleriana realizada, calcula-se a média da intensidade de cada canal em cada quadro, gerando ondas R, G e B. Conforme afirmado em estudos mencionados anteriormente, (ROSA; BETINI, 2020), (SHAO et al., 2016), (BAL, 2015), (LIMA; GUSMÃO; COSTA, 2019), os sinais R e B são utilizados para o cálculo da saturação, enquanto o sinal verde para o cálculo da frequência cardíaca.

Para se estimar a frequência cardíaca, utilizou-se a função da biblioteca *Scipy* (VIRTANEN et al., 2020a) para detectar os picos. Pelos índices dos picos, calcula-se então o intervalo entre um pico e outro pela equação 4.4, sendo FPS a taxa de quadro por segundo, Idx_{pico_i} o índice do i-ésimo pico,$Idx_{pico_{i-1}}$ o índice do pico imediatamente anterior ao i-ésimo pico e $\Delta Pico_i$ o intervalo de tempo entre dois determinados picos. Em seguida é calculado a média destes intervalos, $\Delta Pico_{média}$, sendo N_{picos} o número de picos (Equação 4.5). Por fim, calcula-se os batimentos por minutos, através da equação 4.6.

$$\Delta Pico_i = \frac{Idx_{pico_i} - Idx_{pico_{i-1}}}{FPS} \tag{4.4}$$

$$\Delta Pico_{média} = \frac{\sum_{i=1}^{N_{picos}-1} \Delta Pico_i}{N_{picos} - 1} \tag{4.5}$$

$$BPM = \left\lfloor \frac{60}{\Delta Pico_{média}} \right\rfloor \tag{4.6}$$

Para estimar a saturação de oxigênio, considerou-se o fato de que a oxiemoglobina absorve mais a luz infravermelha, enquanto a hemoglobina absorve mais a luz vermelha. Como a saturação de oxigênio é estimada considerando a relação entre a quantidade de hemoglobina oxigenada, oxiemoglobina, e a quantidade de hemoglobina no sangue, utilizou-se a equação do coeficiente de correlação de Pearson a partir dos sinais vermelho e azul (Equação 4.7).

$$sat = \left| \frac{\sum_{i=0}^{N-1} (V_i - \bar{V})(IV_i - \bar{IV})}{\sqrt{\sum_{i=0}^{N-1} (V_i - \bar{V})^2 \sum_{i=0}^{N-1} (IV_i - \bar{IV})^2}} \right| \tag{4.7}$$

Nesta equação, sat é a saturação,$\bar{V}$, a média do sinal vermelho, V o sinal vermelho, $\bar{IV}$ a média do sinal infravermelho e IV o sinal infravermelho. Quanto maior a saturação, maior será a correlação.

4.2 Algoritmo para Testes Finais

Após os resultados obtidos com os primeiros testes, utilizando-se predominantemente iluminação natural, foi analisados os pontos fortes e fracos do algoritmo inicial, a fim de otimizar os resultados.

4.2.1 Programação em Paralelo

No que se refere à otimização de algoritmos e resolução de problemas, diversas são as alternativas a serem adotadas. Além de otimizar funções e utilizar recursos eficientes, no quesito tempo e eficiência, um recurso bastante eficaz é a programação em paralelo. Esta consiste na utilização dos diversos processadores disponíveis no dispositivo, que realiza o processamento, para a execução das tarefas do algoritmo. Desta forma, múltiplas tarefas são realizadas simultaneamente, de forma síncrona ou assíncrona, possibilitando a redução do tempo, uma vez que as tarefas não são todas executadas em sequência, necessitando da espera do término de uma função para então iniciar a execução de outra. Além da aceleração do programa, a programação em paralelo também possibilita a resolução de problemas complexos, como o processamento de uma grande quantidade de dados, variados e diversos, com fluxo alto. Para a programação em paralelo, podem ser utilizadas *threads* e processos. No algoritmo proposto, foram utilizadas *threads* através da biblioteca *concurrent.futures* (CONCURRENT..., 2024) o que possibilitou a realização de tarefas simultâneas, reduzindo de forma eficaz o tempo de execução, sendo aplicada nas etapas mais lentas do algoritmo.

4.2.2 Aplicação de Métodos Alternativos de Cálculo para Frequência Cardíaca

Outras abordagens podem ser utilizadas para o cálculo da frequência cardíaca, através do sinal obtido pela maximização de vídeo euleriana. Dentre elas, três abordagens alternativas foram testadas na esperança de se conseguir resultados mais próximos. A primeira consistia em identificar os picos do sinal do canal verde, calcular a média $P_{Média}$ da intensidade dos picos (Equação 4.8), e contar o número de picos acima ou igual a média N_{pdm} dividir pelo tempo t do vídeo em segundos e multiplicar por 60, para se obter os batimentos por minuto.

$$P_{Média} = \frac{\sum_{i=1}^{N_{picos}} P_i}{N_{picos}} \tag{4.8}$$

$$BPM = \frac{N_{pdm}}{t} \cdot 60 \tag{4.9}$$

Este método utiliza a média da intensidade dos picos na intenção de atenuar a interferência de ruídos, uma vez que os picos correspondentes ao batimento cardíaco são mais proeminentes que eventuais ruídos. Assim, ao se calcular a média, cria-se um limiar que descarta os picos indesejados.

O segundo método, considerava a hipótese da filtragem do EVM ser suficiente, realizando o mesmo cálculo da equação 4.9, sem a aplicação de nenhum filtro ou janela no sinal verde. Este método serve como um comparativo com os método anteriores, permitindo uma análise prática do efeito da janela de *Hamming* sobre o sinal e do cálculo de média como um limiar.

O terceiro e último método é mais comumente utilizado. Consiste em, a partir do sinal verde, calcular a transformada de Fourier e procurar o maior pico na faixa de frequência correspondente à faixa das possíveis frequências cardíacas. A frequência do maior pico é chamada então de f_{pico}. Multiplica-se a frequência encontrada por 60, para encontrar os batimentos por minutos.

$$BPM = f_{pico} \cdot 60 \tag{4.10}$$

Assim, chamou-se **BPM1**, o método que utiliza a equação 4.9, sem a aplicação da janela e sem o cálculo da média, **BPM2**, quando se utiliza a equação 4.9 com a aplicação da janela e com o cálculo da média como limiar, **BPM3**, quando se utiliza a equação 4.6 com a aplicação da janela sobre o sinal verde e **BPM4**, o método que utiliza a equação 4.10.

4.2.3 Método Alternativo para o Cálculo da Saturação de Oxigênio

Para a saturação de oxigênio, duas novas abordagens foram adicionadas. A primeira abordagem consiste em utilizar o sinal sem ter sido aplicado o filtro adaptativo. Com isto, é possível avaliar as contribuições, sejam positivas ou negativas, do filtro no caso da saturação de oxigênio. Uma vez que o coeficiente de correlação de Pearson avalia a relação entre dois sinais captados pela câmera, a filtragem poderá não estar sendo suficiente. Assim, este foi chamado de **SO1**, enquanto o método utilizado no algoritmo inicial, com o filtro de Kalman, foi chamado de **SO2**.

A segunda abordagem consiste em calcular a taxa R, a partir de Equação 3.5, usada no artigo de Alessandra e Rosa (ROSA; BETINI, 2020). No entanto, este método foi utilizado no estudo mencionado como parte de uma equação linear (Equação 3.6), cujo os coeficientes, α e β, são calibrados para cada voluntário. Por isso, ele apresentará maior eficiência após a aplicação do treinamento realizado no algoritmo.

4.2.4 Espaço de Cor

Continuando o processo de otimização do algoritmo, foram testados também outros espaços de cores. Inspirado pelo texto de Gonzalez (GONZALEZ; WOODS, 2009), optou-se por testar os espaços de cor YCbCr e Lab*, por desacoplarem informações de crominância e luminância, assim como o YIQ, e por serem mais recentes e mais usados atualmente. O primeiro é mais utilizado no sistema de vídeos digital, enquanto o segundo é reconhecido por correlacionar de forma consistente valor da cor e percepção visual da cor. São mais utilizados atualmente e mais eficientes que o YIQ. Assim preserva-se a vantagem do desacoplamento do espaço de cor utilizado anteriormente , adicionando espaços mais eficientes nesta separação de informação de cor.

Também foi testado a variação do fator de amplificação para cada espaço de cor, indo de 1 até 10, buscando o melhor fator de amplificação para cada condição em que o vídeo teria sido gravado. Esta faixa foi escolhida, visando evitar pontos esbranquiçados encontrados no teste anterior, o que poderia afetar a funções de conversão de cor utilizadas.

4.2.5 Transformada de Wavelets Bidimensional

Com a intenção de se obter melhores resultados, optou-se por testar uma abordagem diferente na etapa da magnificação de vídeo Euleriana. Substituiu-se então a pirâmide Gaussiana pela transformada de *Wavelets* Bidimensional com três níveis, assim como era realizado com a pirâmide Gaussiana. Para realização desta, utilizou-se a biblioteca *pywt* (LEE et al., 2019), assim foi possível a aplicação da transformada em cada quadro do vídeo. Para cada nível, obtêm-se um coeficiente de aproximação e outros três coeficientes de detalhes(horizontal, vertical e diagonal). Ao se aplicar mais de um nível na transformada,

o coeficiente de aproximação do nível anterior é substituído pelo resultado da transformada do nível atual, isto é, outros quatro coeficientes: um de aproximação e outros três de detalhes. Considerando três níveis, têm-se um total de dez coeficientes, um de aproximação, três de detalhes horizontais, três de detalhes verticais e três detalhes diagonais. Isso possibilita um processamento mais minucioso e preciso da imagem.

4.2.6 Aprendizado de Máquina

Diversas são as formas de se realizar o aprendizado de máquina, seja desenvolvendo seu próprio treinamento ou ainda utilizando funções prontas fornecidas pelas bibliotecas. Para o algoritmo otimizado, foram testadas quatro abordagens, das quais três consistiam na ideia de se criar um modelo que reduzisse o erro entre o valor estimado e o valor alvo, funcionando como um fator de correção, sendo o primeiro modelo linear e os demais polinomiais, de grau 2 e 3, respectivamente. A quarta abordagem consistia em se utilizar o método de árvore de decisões, que tem o objetivo de estimar os valores alvo a partir do aprendizado de regras de decisões simples, extraídas dos dados apresentados (PEDREGOSA et al., 2011). Basicamente, este último método combina conceitos de predição e árvores de decisão. Para todas as abordagens, foi utilizada a biblioteca *scikit-learn*.

Para analisar e avaliar métodos, o gráfico dos resultados antes do treinamento foi gerado, Figura 8 e Figura 9, onde os pontos estimados flutuavam de forma semelhante aos reais. Logo, regressão linear é um método básico e comum de teste, que apresentaria um custo computacional menor e uma proposta mais simples, valida pro contexto em questão. Regressão polinomial de grau 2 e 3 pode vir a ter maior coerência, uma vez que a relação entre os dados varia ao longo do gráfico, porém terá maior complexidade e custo. Já a árvore de decisão é capaz de encontrar de forma mais precisa e minuciosa a relação entre os dados de entrada e saída, ao dividir esses grupos de dados em subgrupos cada vez menores, criando ramos de decisão suficientes para criar um fator de correção mais adequado.

4.3 Materiais

Todo o projeto foi desenvolvido em *Python* em um computador da Lenovo i7, sétima geração, com placa de video *NVIDIA*. Para as aquisições das imagens, foram utilizadas as câmeras do celular *Samsung A30 S* e do celular *POCO X4 Pro 5G*, sempre apoiadas em um suporte estável, visando minimizar ruídos na imagem. As imagens foram inicialmente capturadas em dois ambientes diferentes, também na tentativa de amenizar as interferências no sinal adquirido. O primeiro ambiente era dentro de uma residência, com as luzes apagadas, pelo período da manhã, entre 9 horas e 10 horas da manhã, iluminado apenas pela luz do sol, que passava pela janela. O segundo ambiente era uma

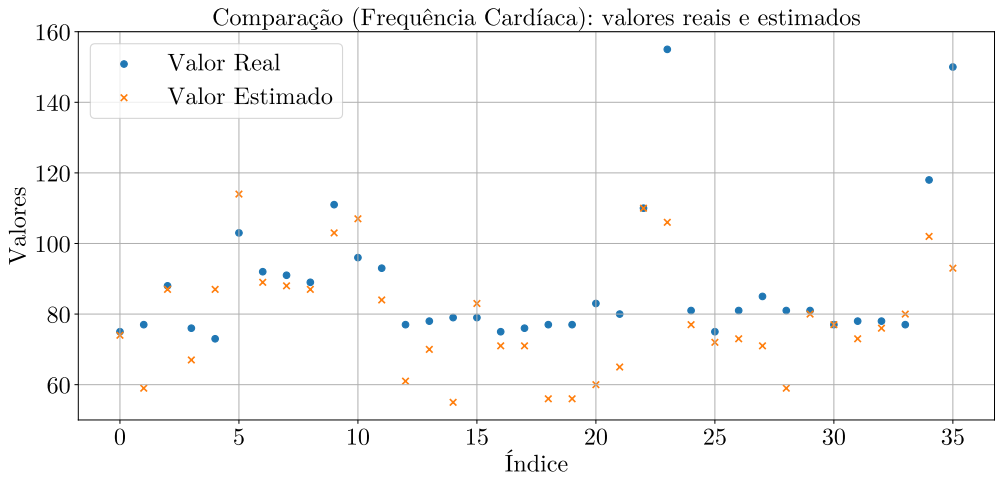


Figura 8 – Gráfico comparativo entre os valores estimados e os valores reais para a frequência cardíaca.

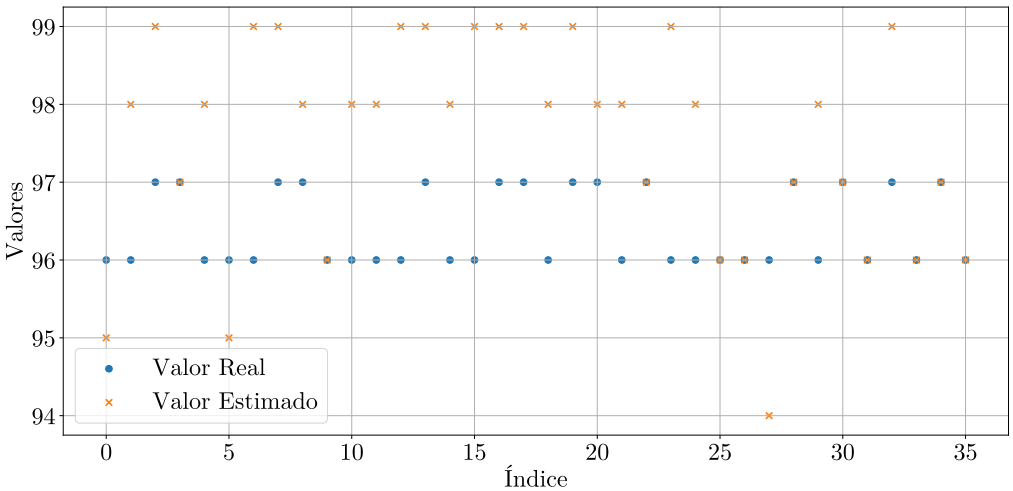


Figura 9 – Gráfico comparativo entre os valores estimados e os valores reais para a saturação de oxigênio.

área externa, sem nenhuma iluminação artificial, com o indivíduo completamente exposto à luz solar. Para a verificação dos resultados foi utilizado um oxímetro de dedo modelo C101H1.

Com os resultados obtidos dos primeiros testes, foi possível obter uma orientação melhorada sobre como aprimorar o algoritmo, a fim de se obter resultados mais coerentes e um software mais robusto. Realizou-se então a gravação de outros trinta vídeos, de 10 segundos cada, entre meia-noite e duas horas da manhã, onde o ambiente apresentava luzes do cômodo apagadas e janelas tampadas por cortinas. A luz dos vídeos era proveniente de um anel iluminador LED, que permitia modificar a coloração da luz entre branca, levemente amarelada, e um tom alaranjado.

5 Resultados e Discussão

5.1 Algoritmo Inicial

Com o algoritmo inicial apresentado no capítulo anterior foi possível obter uma detecção facial satisfatória, em que o indivíduo podia virar o rosto próximo a um ângulo de 90°, em relação à posição em que o rosto fica de frente para a câmera, sem que a detecção falhasse, Figura 10. A detecção escolhida definia *Landmarks* que se mantinham mesmo com estas inclinações. Assim, era possível, através destes pontos, escolher pequenos blocos de interesse no rosto para fazer a medição. A figura 7, é resultado desta detecção facial. Existe, porém, o problema do alto custo computacional desta detecção, resultando em um processamento lento do vídeo (LIMA; GUSMÃO; COSTA, 2019). O custo computacional é relativamente alto e está resultando em um processamento lento do vídeo. Pode-se então utilizar a programação em paralelo para se otimizar este bloco.

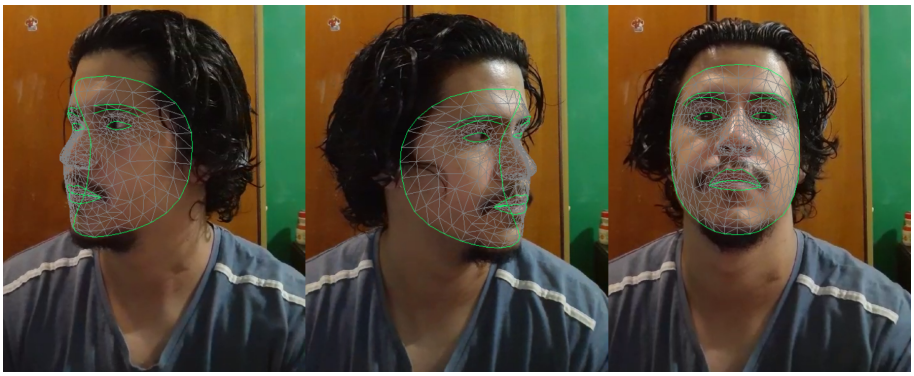


Figura 10 – Detecção facial realizada em diferentes posições do rosto.

A primeira parte do projeto, pode ser considerada bem implementada, alcançando o objetivo desejado de detecção facial e rastreio dos pontos do rosto. A segunda parte se refere à amplificação de vídeo Euleriana. Esta também apresentou resultados eficazes, porém não foram suficientes para se obter os resultados desejados, sendo assim, receberão modificações na segunda parte do desenvolvimento. O método foi eficaz e mostrou-se capaz de realçar as mudanças sutis que ocorrem no rosto, promovidas pelos batimentos cardíacos. Estas mudanças podem ser observadas na figura 11, onde o algoritmo foi aplicado sobre o mesmo video (WU et al., 2012b) utilizado no artigo do Wu (WU et al., 2012a). Uma forma de se aprimorar esta parte do algoritmo pode ser a adição de filtragens ou ferramentas que forneçam a melhor faixa de frequência para a filtragem no tempo, uma vez que foi observado, durante o projeto, que ao se regular as frequências do filtro para uma faixa não tão ampla e mais estreita, próxima da frequência real, melhores eram os resultados. Uma análise mais minuciosa no sinal gerado é capaz de trazer mais informações úteis para

o resultado. No entanto, é preciso preocupar-se com falsos positivos, uma vez que limitar a frequência de corte numa faixa pode ocultar o sinal vital detectado pelos segmentos de vídeo.



Figura 11 – Amplificação de Video Euleriana

Existe, porém, uma vulnerabilidade nesta etapa: luzes artificiais adicionam ruídos ao sinal gerado. A criação de um novo bloco, que lide com este fator, mostra-se promissora, porém a mudança no tratamento da imagem pode ser igualmente eficaz e até mesmo mais prática.

Por fim, para o cálculo dos parâmetros desejados, obteve-se o sinal gerado a partir da média das intensidades das cores em cada quadro. O sinal obtido é apresentado na Figura 12, onde é possível verificar que os sinais dos canais vermelho e azul aparentam apresentar um grau de correlação, o que aponta que a abordagem utilizada para o cálculo da saturação possa ser promissora, ou ao menos útil. Aplica-se sobre estes sinais o filtro de Kalman na expectativa de melhorar a qualidade de sinal, obtendo-se assim o sinal apresentado na Figura 13

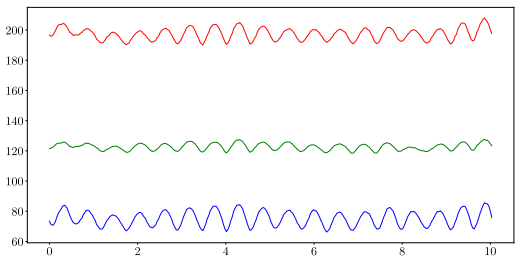


Figura 12 – Sinais obtidos pelas médias das intensidades de cada canal em cada quadro.

Aplicando as equações 4.6 e 4.7, é possível verificar o resultado obtido na figura 14

Em seguida o algoritmo foi aplicado sobre um vídeo desenvolvido pelo próprio autor do presente documento e, assim, foi possível obter os resultados apresentados na Figura 15, no lado esquerdo da imagem. No lado direito da imagem é possível observar o resultado mostrado pelo oxímetro de pulso de dedo. Comparando os dois resultados, é possível

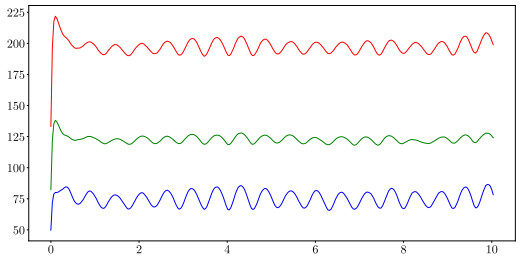


Figura 13 – Sinais filtrados.



Figura 14 – Resultado obtido pela algoritmo

observar que o resultado obtido é próximo ao resultado apresentado pelo oxímetro, na saturação, porém é errôneo na frequência cardíaca, apresentando uma taxa de erro muito alta. Portanto, tornou-se necessário ajustes e mais testes a fim de se verificar possibilidades de soluções e as causas do erro. Uma possibilidade considerada foi a de se alterar o método utilizado para a estimativa do sinal vital por outros que sejam mais assertivos em suas estimativas.



Figura 15 – Resultado obtido pela algoritmo (a esquerda) e estimação apresentada pelo oxímetro de pulso de dedo (a direita)

5.2 Aplicação de melhorias no Algoritmo

5.2.1 Seleção do melhor método de cálculo.

Para melhorar os resultados obtidos anteriormente, foram aplicadas as modificações mencionadas no capítulo anterior, além de se substituir o filtro de Kalman pela janela de Hamming, apenas para o canal verde. Esta, por sua vez, foi aplicada através da função *signal.hamming* da biblioteca *SciPy* (VIRTANEN et al., 2020b). Primeiro foram adicionadas a equação 4.9 e a equação 4.10, como métodos para estimar a frequência cardíaca. Também foi considerado como método adicional para a saturação o cálculo utilizando o coeficiente de correlação de Pearson, porém sobre o sinal sem a aplicação do filtro de Kalman.

Após este acréscimo, o primeiro teste realizado considerou variar unitariamente a amplificação entre 1 e 10, ainda utilizando o espaço de cor **YIQ**. Esta faixa foi escolhida buscando impedir a ocorrência de pixeis com intensidades superiores a 255 após a amplificação, o que resultava em pontos esbranquiçados no vídeo resultante. O teste foi aplicado utilizando-se 6 vídeos de teste, sendo eles: dois em luz branca, dois em luz amarelada e dois em luz laranja. Após testar os 10 valores de amplificação em cada vídeo calculou-se o erro percentual médio de cada amplificação e comparou-se o resultado para se descobrir quais métodos de estimação apresentava o melhor resultado. Conforme exibido na tabela 1 o método **BPM4** e **SO2** apresentaram um erro percentual médio menor que os demais e passaram a ser os métodos selecionados para a estimação nesta etapa de testes.

5.2.2 Seleção do melhor espaço de cor

O segundo teste realizado considera ainda a variação da amplificação, porém, o erro percentual médio calculado considera apenas os pares de vídeos de uma luz específica.

Tabela 1 – Resultado obtido pelo algoritmo variando a amplificação

Amplificação	**BPM1**	**BPM2**	**BPM3**	**BPM4**	**SO1**	**SO2**
Amp-1	328,97	72,9	115,27	11,29	6,74	1,73
Amp-2	310,63	74,21	109,57	22,5	6,75	1,55
Amp-3	286,68	73,04	96,13	22,5	6,75	1,38
Amp-4	252,85	76,48	78,39	22,5	5,02	1,55
Amp-5	226,66	77,55	67,82	23,59	3,28	1,73
Amp-6	205,22	78,85	60,78	27,12	2,24	1,9
Amp-7	194,82	78,85	56,77	29,47	1,9	2,07
Amp-8	175,96	80,13	53,45	29,47	2,07	2,25
Amp-9	164,58	80,13	49,56	29,47	2,25	2,42
Amp-10	149,01	80,13	45,09	29,47	2,25	2,42
Média	229,54	77,23	72,98	**24,74**	3,92	**1,9**

Calcula-se também o desvio padrão para cada caso, resultando em duas tabelas para cada tipo de luz, sendo identificado como **Luz 1** a luz branca, **Luz 2** a luz amarelada e **Luz 3** a luz laranja. Neste ponto, também foram realizados testes com os espaços de cor **L*a*b*** e **YCbCr**, bem como o efeito da amplificação quando aplicados sobre todos os canais, sobre apenas os canais de luminância e, por último, sobre os canais de crominância. De início, os testes são realizados individualmente, resultando em 180 testes por espaço de cor. Desta forma, seis tabelas são geradas, sendo três delas de média e três delas de desvio padrão.

Tabela 2 – Resultado obtido pelo algoritmo variando a amplificação, considerando a média do erro percentual para a Luz Branca, utilizando o espaço de cor L*a*b*.

Amplificação	**BPMlab**	**SOlab**	**BPMl**	**SOl**	**BPMab**	**SOab**
Amp1	23,78	1,56	23,78	1,56	23,78	1,56
Amp2	39,12	1,56	39,12	1,04	23,78	3,12
Amp3	39,12	1,56	39,12	1,56	23,78	4,17
Amp4	39,12	1,04	39,12	2,08	17,87	5,73
Amp5	39,12	1,56	39,12	2,08	17,87	6,77
Amp6	39,12	1,56	39,12	2,60	17,87	7,81
Amp7	39,12	2,08	39,12	2,60	17,87	9,90
Amp8	39,12	2,08	39,12	2,60	17,87	11,46
Amp9	39,12	2,08	39,12	2,60	17,87	13,02
Amp10	39,12	2,60	39,12	2,60	17,87	14,06

A partir das tabelas, procura-se, para cada luz, a amplificação que possua menor erro percentual tanto para os batimentos quanto para a saturação, dando ,de inicio, peso maior para o erro percentual da saturação de oxigênio. Com a seleção de candidatos a melhor fator de amplificação, avalia-se também a tabela com o desvio padrão, fazendo assim a seleção final do valor a ser adotado para cada luz, tratando cada conjunto de canais de forma isolada.

Tabela 3 – Resultado obtido pelo algoritmo variando a amplificação, considerando o Desvio padrão do erro médio percentual para a Luz Branca, utilizando o espaço de cor $L^*a^*b^*$.

Amplificação	**BPMlab**	**SOlab**	**BPMl**	**SOl**	**BPMab**	**SOab**
Amp1	22,45	0,52	22,45	0,52	22,45	0,52
Amp2	7,12	0,52	7,12	1,04	22,45	1,04
Amp3	7,12	0,52	7,12	0,52	22,45	2,08
Amp4	7,12	1,04	7,12	0,00	16,54	4,69
Amp5	7,12	0,52	7,12	0,00	16,54	6,77
Amp6	7,12	0,52	7,12	0,52	16,54	7,81
Amp7	7,12	0,00	7,12	0,52	16,54	8,85
Amp8	7,12	0,00	7,12	0,52	16,54	9,38
Amp9	7,12	0,00	7,12	0,52	16,54	9,90
Amp10	7,12	0,52	7,12	0,52	16,54	10,94

Com a seleção do melhor fator de amplificação para cada conjunto de canais em cada espaço de cor, são comparados os resultados, seguindo o mesmo princípio com relação à média e ao desvio padrão. Assim, considerando a melhor combinação entre menor erro percentual e menor desvio padrão, foi selecionado o espaço de cor YCbCr amplificando os três canais. A partir deste ponto, serão adotadas as amplificações 1, para luz branca, 4, para luz amarela e 3, para a luz laranja.

Tabela 4 – Média do erro percentual.

Espaço(Canal)	**BPM**	**SO2**
lab-lab	17,39	1,55
lab-l	17,39	1,55
lab-ab	17,39	1,04
ycbcr-ycbcr	**11,29**	**1,03**
ycbcr-y	11,29	1,55
ycbcr-cbcr	11,29	1,21
yiq-yiq	11,29	1,73
yiq-y	11,29	1,73
yiq-iq	11,29	1,56

5.2.3 Seleção do Melhor Conjunto de Pontos Faciais

Com as melhores amplificações configuradas e com o melhor espaço de cor selecionado, obteve-se uma melhora de praticamente 18% para a frequência cardíaca, se comparado o erro médio percentual com o erro encontrado pelo primeiro algoritmo. Com o objetivo de melhorar ainda mais o algoritmo, foram comparados os resultados, obtidos utilizando todos os pontos da Figura 7, apenas os pontos da região da testa, pontos da região das bochechas e os pontos de cada bochecha individualmente. Novamente, foram utilizados 6 vídeos para cada conjunto de pontos, resultando em 24 testes realizados. A

Tabela 5 – Desvio padrão do erro percentual.

Espaço(Canal)	**BPM**	**SO2**
lab-lab	14,86	0,79
lab-l	14,86	0,79
lab-ab	14,86	0,85
ycbcr-ycbcr	**7,40**	**0,85**
ycbcr-y	7,40	0,79
ycbcr-cbcr	7,40	0,71
yiq-yiq	7,40	0,48
yiq-y	7,40	0,48
yiq-iq	7,40	0,52

título de identificação, o resultado obtido utilizando todos os pontos foi identificado como **YCbCr - G**, enquanto os demais foram identificados como **YCbCr - T** (testa), **YCbCr - B** (bochecha), **YCbCr - BE** (bochecha esquerda) e **YCbCr - BD** (bochecha direita).

Tabela 6 – Média do erro percentual considerando diferentes pontos faciais.

Teste	**BPM4**	**SO2**
YCbCr - G	**11,29**	**1,03**
YCbCr - T	20,12	1,38
YCbCr - B	25,86	1,73
YCbCr - BE	27,68	2,07
YCbCr - BD	13,50	1,38
Média - YCbCr	19,69	1,52

Com os resultados obtidos, calculou-se novamente a média e o desvio padrão do erro percentual de cada conjunto de pontos. O resultado encontrado indicou que o melhor conjunto era aquele composto tanto pelos pontos da testa, quanto das bochechas, como utilizado anteriormente. Portanto, este conjunto de pontos será adotado para os próximos testes.

Tabela 7 – Desvio padrão do erro percentual considerando diferentes pontos faciais.

Teste	**BPM4**	**SO2**
YCbCr - G	**7,40**	**0,85**
YCbCr - T	10,78	0,49
YCbCr - B	22,43	0,49
YCbCr - BE	20,79	0,01
YCbCr - BD	9,54	0,49
Média - YCbCr	14,19	0,47

5.2.4 Aplicação da transformada de Wavelets Bidimensional

O último teste, antes da aplicação do aprendizado de máquina, consiste em substituir a pirâmide Gaussiana e observar os resultados obtidos. Para este fim, como men-

cionado anteriormente, foi escolhida a transformada de Wavelets bidimensional. Esta transformada produz quatro tipos de coeficientes: coeficientes de aproximação, coeficientes de detalhes horizontais, coeficientes de detalhes verticais e coeficientes de detalhes diagonais. Para a análise de resultados foram utilizados os pontos faciais considerado os pontos da testa e das bochechas. Porém, na parte de desenvolvimento do algoritmo com a transformada em questão, foram utilizados outros seis vídeos, e não os seis mencionados anteriormente, que pertencem aos vídeos que seriam utilizados no treino. Isto se dá pelo fato de que esta parte do algoritmo foi realizada antes dos testes mencionados. Outra diferença é que o espaço de cor utilizado para a configuração do novo modelo, trata-se do L*a*b.

No que se refere aos coeficientes, foram adotados três perspectivas, tanto para a filtragem quanto para a amplificação: Aplicar sobre todos os coeficientes (**GG**), aplicar apenas sobre os coeficientes de aproximação (**GA**) e aplicar apenas sobre os coeficientes de detalhes (**GD**). O primeiro teste das wavelets, comparou a variação dos resultados com relação à filtragem. Conforme exibido na Tabela 8, a melhor combinação entre erro percentual médio e desvio padrão, considerando simultaneamente, frequência cardíaca e saturação de oxigênio, é obtida aplicando a filtragem no tempo apenas sobre os coeficientes de detalhes.

Tabela 8 – Erro percentual considerando diferentes filtragens.

Resultado	**BPMGA**	**BPMGD**	**BPMGG**	**SOGA**	**SOGD**	**SOGG**
Média	21,72	**21,49**	27,41	1,73	**1,90**	1,21
Desvio Padrão	10,71	**10,52**	7,13	1,67	**0,71**	1,11

Este resultado foi obtido considerando uma amplificação aplicada sobre todos os coeficientes. Obtendo-se a melhor abordagem com relação à filtragem, analisa-se, então, qual a melhor abordagem para a amplificação. Desta forma, foi testado, da mesma forma que para a filtragem, a amplificação para os dois tipos de coeficientes e também para cada conjunto de coeficiente de forma isolada. Este teste, conforme mostra a Tabela 9, evidenciou que os melhores resultados são obtidos tanto pela amplificação de todos os coeficientes quanto pela amplificação apenas dos coeficientes de aproximação.

Tabela 9 – Erro percentual considerando diferentes amplificações.

Resultado	**BPMGDA**	**BPMGDD**	**BPMGDG**	**SOGDA**	**SOGDD**	**SOGDG**
Média	**21,49**	24,56	**21,49**	**1,90**	6,59	**1,90**
Desvio Padrão	**10,52**	10,81	**10,52**	**0,71**	8,46	**0,71**

Para verificar se a igualdade entre os resultados não havia sido apenas uma mera coincidência, foram selecionados outros 6 vídeos e realizados novos testes considerando 12 vídeos, porém, comparando apenas os dois melhores resultados. Conforme a Tabela 10, o resultado não se tratou de uma coincidência, mas de um resultado válido. Desta forma,

como os resultados foram iguais, foi escolhida a abordagem com menor custo computacional, isto é, a amplificação aplicada somente sobre os coeficientes de aproximação.

Tabela 10 – Erro percentual considerando amplificações aplicadas sobre todos os coeficientes e sobre apenas os coeficientes de aproximação.

Resultado	**BPMGDA**	**BPMGDG**	**SOGDA**	**SOGDG**
Média	19,15	19,15	2,59	2,59
Desvio Padrão	9,61	9,61	2,65	2,65

Por fim, aplicou-se a variação da amplificação, semelhante a realizada na pirâmide Gaussiana, selecionando como os melhores fatores 1, para luz 1, 5, para luz 2 e 3 para a luz 3. Em seguida, foram testados quatro tipos de Wavelets utilizados na transformada: Daubechies-1,2,3 e 4. Para identificação de cada Daubechies, foi utilizada a sigla db acompanhada do número subscrito correspondente a Wavelet em questão. Inicialmente, para os testes anteriores, foram utilizadas apenas as db_1. Com o fim dos testes, foi verificado que os melhores resultados eram obtidos utilizando a **db_2** e **db_4**. Assim, o critério utilizado foi a seleção do método que apresentou menor custo computacional, isto é, a **db_4**.

Tabela 11 – Erro percentual considerando diferentes tipos de wavelets.

Resultado	**BPM-M**	**BPM-DP**	**SO-M**	**SO-DP**
db_1	17,66	9,20	1,73	0,78
db_2	**15,94**	**10,99**	**1,73**	**1,15**
db_3	21,49	10,52	1,56	1,00
db_4	**15,94**	**10,99**	**1,73**	**1,15**

Com o modelo de EVM com as Wavelets configurado, o penúltimo teste aplicado consiste em revisar os fatores de amplificação, utilizando o espaço de cor YCbCR e os seis vídeos do treino, que estavam sendo utilizados nos testes para a pirâmide Gaussiana. Novamente, então, variou-se a amplificação de 1 até 10 e, com base nos resultados, foram selecionadas as melhores amplificações para a transformada de Wavelets, sendo elas $0,25$ para a luz branca, 5 para a luz amarelada e 3 para a luz laranja. Vale destacar que para luz branca, não foi detectada melhora na faixa de 1 a 10. Portanto, foram testados os valores reduzindo a amplificação em $0,25$ partindo do valor inicial 1. Ao aplicar um fator de $0,25$ o algoritmo apresentou resultados melhores. Porém, para a luz amarelada, apesar da amplificação ter melhorado a partir do fator 2, não foram encontrados resultados melhores, mesmo aumentando o fator até 50. Para isso, optou-se pelo valor médio 5, da faixa de 1 a 10 citada anteriormente, para não manter as amplificações tão baixas.

Com a transformada configurada, comparou-se o resultado obtido pela transformada de Wavelets e pela pirâmide Gaussiana. O resultado, exibido na Tabela 13, considera a seguinte lógica de identificação: O prefixo BPM refere-se a frequência cardíaca

Tabela 12 – Erro percentual considerando a variação da amplificação para a luz 3, Laranja.

Resultado	**BPMMéd**	**BPMDP**	**SOMéd**	**SODP**
Amp_1	9,52	6,95	2,07	0,01
Amp_2	9,52	6,95	2,07	0,01
Amp_3	9,52	6,95	2,07	0,01
Amp_4	23,62	7,15	2,07	0,01
Amp_5	27,15	3,62	2,07	0,01
Amp_6	27,15	3,62	2,07	0,01
Amp_7	27,15	3,62	2,07	0,01
Amp_8	27,15	3,62	2,07	0,01
Amp_9	27,15	3,62	2,07	0,01
Amp_10	27,15	3,62	2,07	0,01

e o prefixo SO refere-se a saturação de oxigênio. Estes são sucedidos pela letra W, para as transformadas de Wavelets, ou G, para a Pirâmide Gaussiana. Apesar da saturação possuir resultados melhores utilizando-se a pirâmide Gaussiana, as Wavelets apresentaram resultados consideravelmente melhores para a frequência cardíaca, portanto serão adotadas como método final, uma vez que é possível compensar o erro da saturação com o aprendizado de máquina planejado.

Tabela 13 – Erro percentual considerando a aplicação do EVM com as pirâmides Gaussianas e com a transformada de Wavelets.

Resultado	**BPMG**	**BPMW**	**SOG**	**SOW**
Média	11,29	**7,81**	1,03	**1,90**
Desvio Padrão	7,40	**5,09**	0,85	**0,38**

5.2.5 Aplicação do aprendizado de máquina.

Com a configuração do processamento de imagem ajustado, a próxima etapa seria a realização do treino do algoritmo para se implementar o aprendizado de máquina, com o objetivo de se criar um fator de correção que ajuste o resultado obtido para valores mais próximos do real. Para isso, foram gravados doze vídeos testes, os quais, seis eram filmados com o usuário em repouso e os outros seis, depois de um curto período de atividade física de alta intensidade, com o objetivo de estimular a frequência cardíaca. Com o término das filmagens, identificou-se que a faixa escolhida para a filtragem, $0,833Hz$ a $2Hz$, não englobaria as altas frequência obtidas. Optou-se então por ampliar a faixa para $3Hz$, a mesma utilizada no artigo (KOLOSOV et al., 2023). Foram realizados quatro treinamentos a partir de trinta vídeos, disponível em: (TRINDADE, 2024), gravados com o usuário em repouso, sendo destes trinta, dez para cada tipo de luz. Os três primeiros treinos eram dedicados, respectivamente, para cada tipo de luz, sendo o primeiro treino para luz branca, o segundo para luz amarela e o terceiro para a luz laranja, sendo cada

um deles alimentados com dez vídeos. Para cada um, quatro modelos foram utilizados, como mencionados anteriormente. O quarto treino foi realizado com os trinta vídeos e também possui quatro modelos. Após a realização do treinamento, os quatro modelos eram testados com os doze vídeos e calculava-se então o erro percentual médio e o desvio padrão deste erro. Vale ressaltar que os doze vídeos em questão, 345 quadros, não fazem parte do conjunto de trinta vídeos do treino, 781 quadros. Selecionava-se então um limiar para o erro médio, este limiar, para as frequências cardíacas, foi selecionado iniciando-o como 20% e reduzindo unitariamente a cada teste, até não ser encontrado nenhum modelo com erros abaixo do limiar. Com o conjunto de modelos selecionados, procurava-se o modelo com menor desvio padrão. O mesmo procedimento foi adotado para a saturação de oxigênio, sendo o limiar inicial selecionado $0,5\%$, reduzindo em $0,1\%$ a cada teste.

Na etapa de treinamento, utilizou-se também o cálculo da taxa R citado por Rosa *et al.* (ROSA; BETINI, 2020) para o cálculo da saturação de oxigênio, além do método utilizado anteriormente. Com estas configurações, o resultado encontrado é apresentado na Tabela 14, onde é possível observar que o treinamento que apresentou melhor resultado foi aquele que utilizou os trinta vídeos, apresentando menor erros para as três luzes usadas.

Na Tabela 14, as siglas apresentam o seguinte significado:

- M1 - Regressão Linear
- M2 - Regressão Polinomial de grau 2
- M3 - Regressão Polinomial de grau 3
- M4 - Árvore de decisões
- S1 - Saturação calculada pelo coeficiente de correlação de Pearson utilizando os sinais azul e vermelho, após a aplicação do filtro de Kalman.
- S2 - Saturação calculada a partir da taxa R mencionada por Rosa *et al.* (ROSA; BETINI, 2020)

Tabela 14 – Erro médio percentual encontrado ao ser aplicado os quatro tipos de treino.

Treino	**Média (BPM/Sat)**	**Desvio Padrão (BPM/Sat)**	**Modelo-BPM**	**Modelo-Sat.**	**Met.-Sat**
Treino - Luz 1	6% / 0,4%	2 / 0,1	M4	M2	S1
Treino - Luz 2	15,6% / 0,5%	9,8 / 0,1	M2	M3	S2
Treino - Luz 3	19,6% / 0,3%	11,2 / 0,4	M4	M4	S1
Treino - Geral	**6,6% / 0,4%**	**4,9 / 0,4**	**M4**	**M4**	**S1**

Em seguida, adicionou-se seis vídeos, com os batimentos cardíacos acelerados, ao treinamento e realizou-se outro treino considerando os trinta e seis vídeos, 958 quadros, e os quatro modelos. Com isto, o resultado obtido é apresentado na Tabela 15, onde é possível observar uma melhora nos resultados.

Tabela 15 – Erro médio percentual encontrado ao ser aplicado o treino geral com os trinta e seis vídeos.

Treino	**Média (BPM/Sat)**	**Desvio Padrão (BPM/Sat)**	**Modelo-BPM**	**Modelo-Sat.**	**Met.-Sat**
Treino - Geral	3,9% / 0,3%	5,6 / 0,3	M4	M4	S2

Aplicando o algoritmo no vídeo da Figura 15, encontra-se uma melhora significativa, apresentando um erro de apenas 4 BPM. Um erro bem menor do que o observado anteriormente, conforme exibido na Figura 16.



Figura 16 – Resultado obtido pela novo algoritmo proposto

Os novos resultados parecem promissores, considerando o mesmo usuário, conseguindo bons resultados para diferentes tipos de iluminação e para diferentes tipos de faixas de frequência. Possivelmente, aumentar o número de vídeos para o treinamento, acrescentar mais casos de frequência, abaixo e acima do normal, e aumentar o número de usuários tornaria o algoritmo muito mais robusto e resistente a variações. No geral, o programa funciona bem para a faixa normal do usuário, tanto sem o aprendizado de máquina, quanto com o mesmo, apresentando maiores oscilações apenas nas frequência muito altas e muito baixas, o que é reduzido pelo treinamento. Como pode ser observado na Figura 17.

5.2.6 Testando o efeito do resultado para diferentes taxas de quadro por segundo.

Após a conclusão do algoritmo e a obtenção dos resultados, foram selecionados dois vídeos, o vídeo "**face.mp4**"(WU et al., 2012b) e um novo vídeo gravado ao meio dia dentro de uma residência com a luz acesa. Variou-se a taxa de quadros destes vídeos de 30 *fps* até

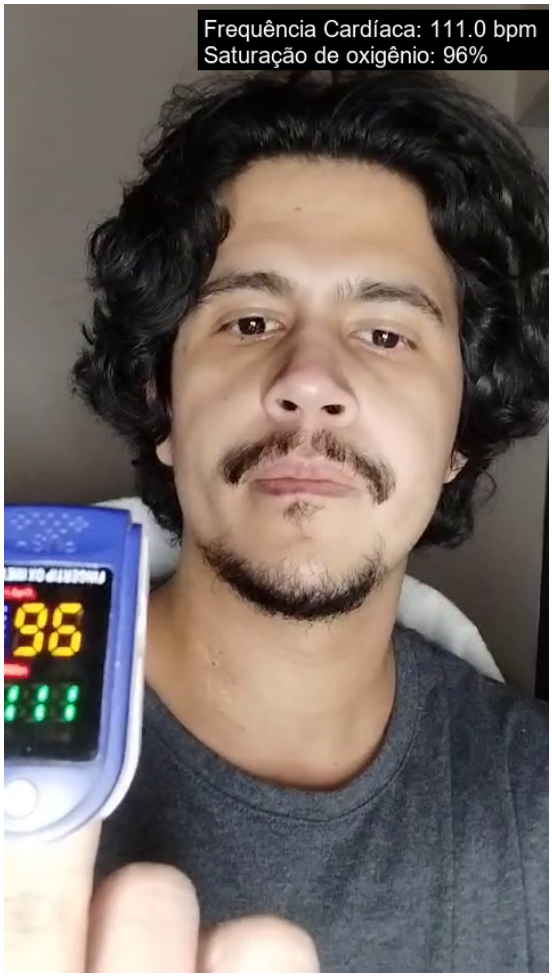


Figura 17 – Resultado obtido pela novo algoritmo proposto

10 *fps*, sempre reduzindo em 5 *fps*, obtendo-se outros oito vídeos, quatro para cada vídeo original. Com isso, foi possível verificar a robustez do programa proposto para vídeos com diferentes taxas de transmissão. Observou-se, porém, como o treinamento foi aplicado somente sobre vídeos do usuário, que o aprendizado de máquina forçava uma correção errônea para baixas frequências, uma vez que nos vídeos do *dataset* nunca ocorreram frequências nessa faixa. Por esta razão, foi desativado o módulo de correção para o vídeo "**face.mp4**". Com isto, o resultado encontrado é apresentado na Tabela 16, onde é possível ver que o resultado se manteve estável e próximo ao real para diferentes taxas, o que possibilita não só a redução de taxas para um processamento mais rápido, mas também uma resistência maior a vídeos com taxas de quadro menores. No entanto, considerando os resultados obtidos por Lima *et al.* em seu artigo(LIMA; GUSMÃO; COSTA, 2019), recomenda-se, para análise de frequências cardíacas médias e instantâneas, limitar a 15 *fps* a taxa de quadros mínima.

Tabela 16 – Resultado obtido com a variação da taxa de quadros por segundo

Vídeo	10 FPS	15 FPS	20 FPS	25 FPS	30 FPS	Valor Real
Vídeo próprio(BPM)	79	79	79	79	79	75
Vídeo próprio(SO2)	96%	97%	97%	96%	96%	97%
Vídeo próprio(BPM)	53	54	54	54	54	53

5.2.7 Testes de Validação

Para a validação do algoritmo, optou-se por realizar o cálculo do *F1 score*, que permite uma avaliação simultânea das métricas de precisão e *recall* através da média harmônica dos dois valores (PEDREGOSA et al., 2011). Para isto, é necessário ter, antes, o cálculo das duas métricas.

A precisão irá indicar o grau de confiança de um modelo, isto é, o quanto ele acerta na sua classificação de um caso como positivo em relação a quantos ele previu como positivo. A equação utilizada para o cálculo de precisão é dado pela Equação 5.1.

$$Pr = \frac{VP}{VP + FP} \tag{5.1}$$

O *recall* indica quantas classificações positivas o modelo acertou em relação a todos os reais positivos, ou seja, de todos os casos positivos, quantos ele teria acertado. Esta métrica também é chamada de taxa de detecção e pode ser calculada pela Equação 5.2.

$$Rc = \frac{VP}{VP + FN} \tag{5.2}$$

Com os dois valores encontrados, o *F1 score* pode ser estimado com a Equação 5.3.

$$F1_{score} = 2 \cdot \frac{Pr \cdot Rc}{Pr + Rc} \tag{5.3}$$

No entanto, antes da realização destes cálculos, é necessário estabelecer o conceito de verdadeiro positivo (VP), verdadeiro negativo (VN), falso positivo (FP) e falso negativo (FN) no contexto do trabalho em questão. Na realidade do projeto abordado, a classificação terá o seguinte significado:

- **Verdadeiro Positivo(VP)**: Valores dentro da faixa de erro do oxímetro que sejam maiores ou igual ao valor apresentado pelo oxímetro.
- **Verdadeiro Negativo(VN)**: Valores dentro da faixa de erro do oxímetro que sejam menores do que o valor apresentado pelo oxímetro.
- **Falso Positivo(FP)**: Valores fora da faixa de erro do oxímetro que sejam maiores que o valor apresentado pelo oxímetro.

- **Falso Negativo(FN)**: Valores fora da faixa de erro do oxímetro que sejam menores do que o valor apresentado pelo oxímetro.

Para que estes critérios possam ser considerados, é necessário ter a precisão do oxímetro, no entanto, o documento físico do dispositivo foi descartado em 2021, e não foi encontrado uma versão digital do modelo referente. Para resolver este impasse, foi adotado como referência os valores mencionados no manual do oxímetro *Premium* da **DellaMED** (HOSPITALAR, 2017) que apontam para uma acurácia de ±2% para a saturação de oxigênio, de ±2 BPM, para frequências cardíacas entre 30 e 99 BPM, e ±2% para frequências entre 100 e 250 BPM. Estes valores são coerentes com os mencionados no artigo de Rosa *et al.* (ROSA; BETINI, 2020).

Com a referência definida, o algoritmo foi testado em 31 vídeos, não pertencentes ao treinamento, que haviam sido gravados, em parte, em excesso, durante a gravação do treino e do teste. O conjunto de gravações contém vídeos em iluminação artificial branca, amarela, laranja e também com a iluminação do ambiente. Aproveitando este conjunto de vídeos, conjunto geral, foi calculado as mesmas métricas considerando apenas os vídeos com luz branca, conjunto com luz branca, amarela, conjunto com luz amarela, e laranja, conjunto com luz laranja, separadamente.

Tabela 17 – Resultado do F1 score para os diferentes conjuntos de vídeo.

Conjunto de vídeo	**$F1_{score}$(BPM)**	**$F1_{score}$(sat.)**
Conjunto geral	0,71	1,0
Conjunto com luz branca	0,72	1,0
Conjunto com luz amarela	0,44	1,0
Conjunto com luz laranja	0,94	1,0

É possível observar que, tanto para o conjunto geral, quanto para o conjunto com luz branca e laranja, o *F1* score para a frequência cardíaca foi alta, sendo maior para a luz laranja, indicando que o treinamento foi realizado com qualidade, tendo bons resultados. No entanto, para luz amarela, o algoritmo proposto apresentou uma limitação maior, se tornando não recomendado para este tipo de iluminação.

Por outro lado, a saturação de oxigênio teve *score* máximo para todos os casos. Isto se dá pela limitação de dados fornecidos para o modelo, uma vez que o usuário em questão não conseguiu afetar sua saturação para obter valores fora da faixa entre 96% e 97%. Mesmo segurando a respiração até sentir desconforto, não eram detectadas variações na sua saturação fora desta faixa. Portanto, apesar de parecerem ótimos resultados, podem apontar para uma limitação, sendo errôneo para baixas saturações. Mesmo assim, o resultado indica que dentro destes limites, a saturação estimada tende a ter resultados válidos.

Foi avaliado, também, a similaridade entre o conjunto de valores estimados e o conjunto de valores reais, por meio do Índice de Jaccard, que pode ser calculado pela equação 5.4. No entanto, no contexto em questão, considerando os verdadeiros positivos e negativos, bem como os falsos negativos e positivos e seus significados, o índice pode ser calculado conforme exibido na equação 5.5

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{5.4}$$

$$IoU = \frac{VP}{VP + FN + FP} \tag{5.5}$$

Aplicando estas equações nos testes realizados com os trinta e um vídeos, mencionados anteriormente, e considerando os conjuntos de vídeos usados no teste anterior, obteve-se os resultados apresentados na Tabela 18.

Tabela 18 – Resultado do IoU para os diferentes conjuntos de vídeo.

Conjunto de vídeo	**IoU(BPM)**	**IoU(sat.)**
Conjunto geral	0,55	1,00
Conjunto com luz branca	0,57	1,00
Conjunto com luz amarela	0,29	1,00
Conjunto com luz laranja	0,89	1,00

O resultado indica, novamente, uma similaridade perfeita para a saturação provocada pela baixa variação dos dados. Porém, indica um comportamento coerente para a faixa de 96% até 97%.

Por outro lado, considerando a frequência cardíaca, os resultados são positivos para todos os conjuntos, exceto para os vídeos com luz amarela, reforçando a sensibilidade do algoritmo a esta condição. No entanto, nas demais situações, a similaridade é boa e promissora, porém ainda baixa, para o conjunto geral e com luz branca. O que indica que, para estes casos, ainda cabem melhorias possíveis. Porém o trabalho apresenta mais acertos que erros.

6 CONCLUSÃO

Diante da experiência vivenciada pelo mundo em 2020 com a pandemia provocada pelo vírus *SARS-CoV-2*, potencialmente grave e de elevada taxa de transmissibilidade, ficou evidente a necessidade da realização de certas tarefas de forma remotas, incluindo o atendimento a pacientes, que por vezes ficaram afastados de suas rotinas de acompanhamento pelo risco de infecção. Somada a isso, existe contemporaneamente a busca por métodos não-invasivos e de maior conforto ao usuário, o que possibilita uma análise contínua e prolongada do indivíduo avaliado.

Dentro deste cenário, o trabalho em questão apresentou uma proposta que permite aferir a frequência cardíaca e a saturação de oxigênio sem a necessidade de nenhum tipo de contato com o corpo, utilizando apenas segmentos de vídeo. O algoritmo em questão foi capaz de realizar a detecção facial, para extrair os pontos da região de interesse com eficiência, além de realizar com sucesso o processo de magnificação de vídeo Euleriana, sugerido por Wu *et al.*(WU et al., 2012a), substituindo apenas a pirâmide Gaussiana pela transformada de Wavelets, o que trouxe melhorias significativas no resultado obtido. A transformada de Wavelets mostrou-se bastante promissora por permitir uma análise mais detalhada e precisa das variações nos quadros.

Por fim, com a imagem tradada, realizou-se a filtragem do sinal no tempo e obteve informações sobre a frequência cardíaca, através do canal verde, e sobre a saturação de oxigênio, por meio dos canais azul e vermelho. Como o sinal se encontrava ruidoso, foi necessário aplicar o filtro de Kalman sobre o sinal dos canais vermelho e azul, e a janela de hamming sobre o canal verde para melhorar os resultados. Com o sinal filtrado, aplicou-se a FFT sobre o canal verde para se obter a frequência cardíaca, conseguindo resultados próximos, falhando apenas em alguns casos, onde o erro apresentava um valor percentual maior. No caso da saturação de oxigênio, o coeficiente de correlação de Pearson foi bastante eficiente, porém acabou apresentando resultados piores que os obtidos pela taxa R, utilizada por Rosa *et al.*(ROSA; BETINI, 2020), quando o treinamento foi aplicado utilizando os trita e seis vídeos.

O aprendizado de máquina, por sua vez, trouxe resultados promissores, porém limitados ao usuário. Um acréscimo de vídeos e indivíduos analisados pode resultar num aperfeiçoamento ainda mais eficiente, reduzindo o risco da não detecção de alguma arritmia. No entanto, dada as condições de análise e a comparação com o oxímetro, a taxa de erro médio percentual de $3,9\%$ para os batimentos cardíacos e $0,3\%$ para a saturação de oxigênio apontam para ótimos resultados, atendendo aos objetivos do trabalho.

O teste de validação ainda mostrou um *F1 score* alto para os conjuntos com todos

os vídeos, com luz branca e com luz laranja, sendo este último o com resultados mais promissores e precisos, o que indica uma confiança alta para o algoritmo nesta condição. Porém, nestas condições, o trabalho se apresentou similar ao oxímetro apenas para a luz laranja, o que indica que a luz branca pode apresentar maiores erros de medições. Um número maior de vídeos aplicados no processo de treinamento, com uma variação da frequência cardíaca maior, poderá aprimorar os resultados obtidos, mas, o software voltado para a câmera do celular, acrescenta ruídos a imagem ao tentar ajustar a qualidade do vídeo de forma autônoma. O uso de câmeras que não tenha esse processamento automático possibilitaria uma análise mais próxima da encontrada pela referência.

No caso da luz amarela, os resultados foram bem abaixo do esperado, tendo uma oscilação maior nos resultados. Boa parte desta dificuldade pode se dar pelo próprio algoritmo da câmera, mas seria necessário uma análise mais profunda, embora o acréscimo de robustez no treino pudesse otimizar os resultados também. Outras opções viáveis seriam considerar um treino específico para a luz amarela, utilizando este modelo para estas condições, seleção de outros tipos de filtro, adição de módulos de normalização da iluminação, ou ainda adição de módulos adicionais de estabilização da imagem.

Já para a saturação, o resultado se mostrou confiável para a faixa de 96% até 97%. Seria necessário uma amostra maior de usuários para verificar a eficiência do projeto, uma vez que a faixa estava limitada ao usuário que apresentou uma variação muito baixa na saturação. Porém, considerando os recursos disponíveis, o algoritmo se mostrou bastante promissor e superou as expectativas, porém ainda provoca receios em virtude da faixa restrita.

O algoritmo também apresentou uma robustez elevada considerando uma variação de quadros por segundo, possibilitando mais uma otimização. O erro, manteve-se constante e baixo entre a faixa de 30 até 10 *fps*, chegando a reduzir a 15 *fps* para a saturação de oxigênio. O tempo de execução reduz significativamente, mantendo a qualidade do resultado, a taxas menores de transmissão, o que possibilita o acréscimo de mais um módulo, caso haja a necessidade.

Apesar dos resultados positivos quanto às medições, o algoritmo permanece com um custo computacional alto e limitado a variações de iluminação intensa durante o vídeo, isto é, alternâncias bruscas entre claro, alta iluminação, e escura, baixa ou nenhuma iluminação. Quanto ao custo computacional, formas de se aprimorar o algoritmo seria a otimização do código compilado em Python, uso de bibliotecas otimizadas ou ainda mudança de linguagem.

Assim, o trabalho atende aos objetivos desejados, sendo possível ainda trazer melhorias por meios de modificações de hardware, para a otimização do tempo de execução, acréscimos de vídeos ao *dataset* e uso de câmeras não sujeitas a otimizações automáticas. Outra melhoria possível de ser aplicar é utilizar técnicas de análise de componentes

independentes (ICA) para separar o ruído do sinal cardíaco do canal verde, ou ainda, como módulo complementar da filtragem temporal aplicado sobre os quadros do vídeo, buscando separa as informações relevantes de ruídos, trazendo assim resultados com erros bem menores (HYVäRINEN; OJA, 2000).

Referências

ALBERTO, C.; PEREIRA, C.; MOREIRA, M. Â. F. Pletismografia–resistência das vias aéreas. *J Pneumol*, v. 28, n. Suppl 3, p. 139, 2002. Citado na página 33.

BAL, U. Non-contact estimation of heart rate and oxygen saturation using ambient light. *Biomed. Opt. Express*, Optica Publishing Group, v. 6, n. 1, p. 86–97, Jan 2015. Disponível em: <https://opg.optica.org/boe/abstract.cfm?URI=boe-6-1-86>. Citado 2 vezes nas páginas 37 e 42.

COBOS-TORRES, J.-C.; ABDERRAHIM, M. Simple measurement of pulse oximetry using a standard color camera. In: *2017 40th International Conference on Telecommunications and Signal Processing (TSP)*. [S.l.: s.n.], 2017. p. 452–455. Citado na página 24.

CONCURRENT.FUTURES — Launching parallel tasks. 2024. <https://docs.python.org/3/library/concurrent.futures.html>. Citado na página 43.

GONZALEZ, R.; WOODS, R. *Processamento digital de imagens (3a. ed.)*. Pearson Educación, 2009. ISBN 9781512939958. Disponível em: <https://books.google.com.br/books?id=UgtMswEACAAJ>. Citado 2 vezes nas páginas 39 e 45.

HOSPITALAR, D. *Oxímetro de pulso digital MANUAL DO USUÁRIO (Português)*. 2017. Disponível em: <https://magazinemedica.com.br/media/images/ProductFile/26a116c3b7c278198ee2cc074fb98a3f.pdf?srsltid=AfmBOoprGad19LnQzZjY6DPTMZb_4cPFOLrrpRakocrXb8p-sNbQZGoa>. Citado na página 63.

HYVäRINEN, A.; OJA, E. Independent component analysis: algorithms and applications. *Neural Networks*, v. 13, n. 4, p. 411–430, 2000. ISSN 0893-6080. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0893608000000265>. Citado na página 67.

KOLOSOV, D. et al. Contactless camera-based heart rate and respiratory rate monitoring using ai on hardware. *Sensors*, v. 23, n. 9, 2023. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/23/9/4550>. Citado na página 58.

LEE, G. R. et al. Pywavelets: A python package for wavelet analysis. *Journal of Open Source Software*, The Open Journal, v. 4, n. 36, p. 1237, 2019. Disponível em: <https://doi.org/10.21105/joss.01237>. Citado na página 45.

LIMA, A. T. M.; GUSMÃO, D. B.; COSTA, M. V. C. Remote detection of heart beat and heart rate from video sequences. In: COSTA-FELIX, R.; MACHADO, J. C.; ALVARENGA, A. V. (Ed.). *XXVI Brazilian Congress on Biomedical Engineering*. Singapore: Springer Singapore, 2019. p. 437–440. ISBN 978-981-13-2517-5. Citado 5 vezes nas páginas 24, 35, 42, 49 e 61.

LLC, G. *MediaPipe: A Framework for Building Perception Pipelines*. 2022. <https://mediapipe.dev/>. Accessed: 28 de setembro de 2024. Citado 2 vezes nas páginas 36 e 41.

LUDERMIR, T. B. Inteligência artificial e aprendizado de máquina: estado atual e tendências. *Estudos Avançados*, Instituto de Estudos Avançados da Universidade de São Paulo, v. 35, n. 101, p. 85–94, Jan 2021. ISSN 0103-4014. Disponível em: <https://doi.org/10.1590/s0103-4014.2021.35101.007>. Citado na página 39.

LYNCH, P. J. *Heart left anterior oblique xs*. 2006. Description: Heart left anterior oblique cross section internal anatomy. Disponível em: <https://creativecommons.org/licenses/by/2.5/?ref=openverse>. Citado na página 27.

MADHERO88. *Electrical conduction system of the heart*. 2009. Disponível em: <https://creativecommons.org/licenses/by-sa/3.0/?ref=openverse>. Citado na página 29.

MEJIA-MEJIA, E. et al. Photoplethysmography signal processing and synthesis. *Photoplethysmography: Technology, Signal Analysis, and Applications, PA Kyriacou and J. Allen, Eds. London, UK: Elsevier*, p. 69–145, 2021. Citado na página 34.

NOWARA, E. M.; MCDUFF, D.; VEERARAGHAVAN, A. Combining magnification and measurement for non-contact cardiac monitoring. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. [S.l.: s.n.], 2021. p. 3805–3814. Citado na página 24.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011. Citado 2 vezes nas páginas 46 e 62.

PETERHCHARLTON. *Photoplethysmogram (PPG) signal*. 2020. Photoplethysmogram (PPG) signal, detected onsets (red circles) and a labelled inter-beat interval. Disponível em: <https://creativecommons.org/licenses/by/4.0/?ref=openverse>. Citado na página 34.

PETERHCHARLTON. *Photoplethysmogram signal components*. 2021. This figure illustrates the simplified origins of light attenuation which form the PPG signal (arterial blood, venous blood and other tissues). The figure is adapted from slide 7 of: Charlton, P. H. (2018, August). Capitalising on Smart Wearables to Improve Health Monitoring [Data set]. Zenodo. https://doi.org/10.5281/zenodo.1406011. The figure was created using data from the 'BIDMC PPG and Respiration' dataset: https://doi.org/10.13026/C2208R. Disponível em: <https://creativecommons.org/licenses/by/4.0/?ref=openverse>. Citado na página 34.

POH, M.-Z.; POH, Y. C. Validation of a standalone smartphone application for measuring heart rate using imaging photoplethysmography. *Telemedicine and e-Health*, v. 23, n. 8, p. 678–683, 2017. PMID: 28140834. Disponível em: <https://doi.org/10.1089/tmj.2016.0230>. Citado na página 24.

ROSA, A. d. F. G.; BETINI, R. C. Monitoramento da taxa de saturação de oxigênio no sangue e frequência cardíaca via método de magnificação euleriana sem contato físico. *Revista Brasileira de Computação Aplicada*, v. 12, n. 2, p. 80–92, 2020. Citado 8 vezes nas páginas 24, 37, 38, 42, 45, 59, 63 e 65.

SANTOS, N. *Anatomia e Fisiologia Humana*. [s.n.], 2018. ISBN 9788536507170. Disponível em: <https://books.google.com.br/books?id=H4diDwAAQBAJ>. Citado na página 27.

SCULLY, C. G. et al. Physiological parameter monitoring from optical recordings with a mobile phone. *IEEE Transactions on Biomedical Engineering*, v. 59, n. 2, p. 303–306, 2012. Citado na página 37.

SHAO, D. et al. Noncontact monitoring of blood oxygen saturation using camera and dual-wavelength imaging system. *IEEE Transactions on Biomedical Engineering*, v. 63, n. 6, p. 1091–1098, 2016. Citado na página 42.

SHI, J.; TOMASI, C. Good features to track. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 593–600, 1994. Citado na página 35.

SÖRNMO, L.; LAGUNA, P. *Bioelectrical signal processing in cardiac and neurological applications*. [S.l.]: Academic press, 2005. v. 8. Citado 4 vezes nas páginas 23, 28, 30 e 32.

TARASSENKO, L. et al. Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. *Physiological Measurement*, v. 35, p. 807, 03 2014. Citado na página 37.

TRINDADE, R. R. L. *Estimação de frequência cardíaca e saturação de oxigênio a partir de segmentos de vídeo usando técnicas de processamento de imagem*. 2024. Disponível em: <https://github.com/RRLTrinity/TCC_2>. Citado na página 58.

TUSMAN, G.; BOHM, S. H.; SUAREZ-SIPMANN, F. Advanced uses of pulse oximetry for monitoring mechanically ventilated patients. *Anesthesia & Analgesia*, Wolters Kluwer, v. 124, n. 1, p. 62–71, 2017. Citado 2 vezes nas páginas 24 e 33.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, p. I–511 – I–518, 2001. Citado na página 35.

VIRTANEN, P. et al. Scipy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, v. 17, n. 3, p. 261–272, 2020. Disponível em: <[1](https://scipy.org/citing-scipy/)>. Citado na página 42.

VIRTANEN, P. et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, v. 17, p. 261–272, 2020. Citado na página 52.

WALT, S. Van der et al. scikit-image: image processing in python. *PeerJ*, PeerJ Inc., v. 2, p. e453, 2014. Citado na página 41.

WU, H.-Y. et al. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (Proc. SIGGRAPH 2012)*, v. 31, n. 4, 2012. Citado 4 vezes nas páginas 24, 36, 49 e 65.

WU, H.-Y. et al. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics (Proc. SIGGRAPH 2012)*, v. 31, n. 4, 2012. Disponível em: <[1](https://people.csail.mit.edu/mrub/evm/)>. Citado 2 vezes nas páginas 49 e 60.

YITZHAKNAT. *Heart rate variability (HRV)*. 2022. English: Heart rate variability with basic ECG waves, intervals and segments (edited). Disponível em: <https://creativecommons.org/licenses/by-sa/4.0/?ref=openverse>. Citado 3 vezes nas páginas 11, 31 e 32.