

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

**Correção de cor para tons de pele em imagens adquiridas
por dispositivos móveis**

Relatório científico final de iniciação científica referente ao processo FAPESP nº 2023/15404-1.

Bolsista: Rodrigo Rodrigues de Castro

Orientador: João do Espírito Santo Batista Neto

São Carlos, SP

2024

Sumário

1	Introdução	.	.	.	3
1.1	Projeto e cronograma originais	.	.	.	3
2	Atividades realizadas	.	.	.	5
2.1	Filtragem Bilateral	.	.	.	5
2.2	Integração dos Algoritmos	.	.	.	7
3	Resultados obtidos	.	.	.	8
4	Considerações Finais	.	.	.	13
Referências					14

1 Introdução

Este documento tem como propósito relatar as atividades realizadas pelo bolsista Rodrigo Rodrigues de Castro no período de julho de 2024 a dezembro de 2024, referentes ao projeto de Iniciação Científica, processo FAPESP nº 2023/15404-1, intitulado "Correção de cor para tons de pele em imagens adquiridas por dispositivos móveis". O projeto visa o desenvolvimento de um método de correção de cor de pele humana (*whitebalance*) a partir de uma imagem de referência.

1.1 Projeto e cronograma originais

Este projeto de pesquisa tem como finalidade principal o desenvolvimento de um método de correção de cor de pele da face humana (*whitebalance*) a partir de imagens adquiridas por dispositivos móveis, que podem apresentar uma variabilidade considerável de tonalidade entre amostras de uma mesma pessoa.

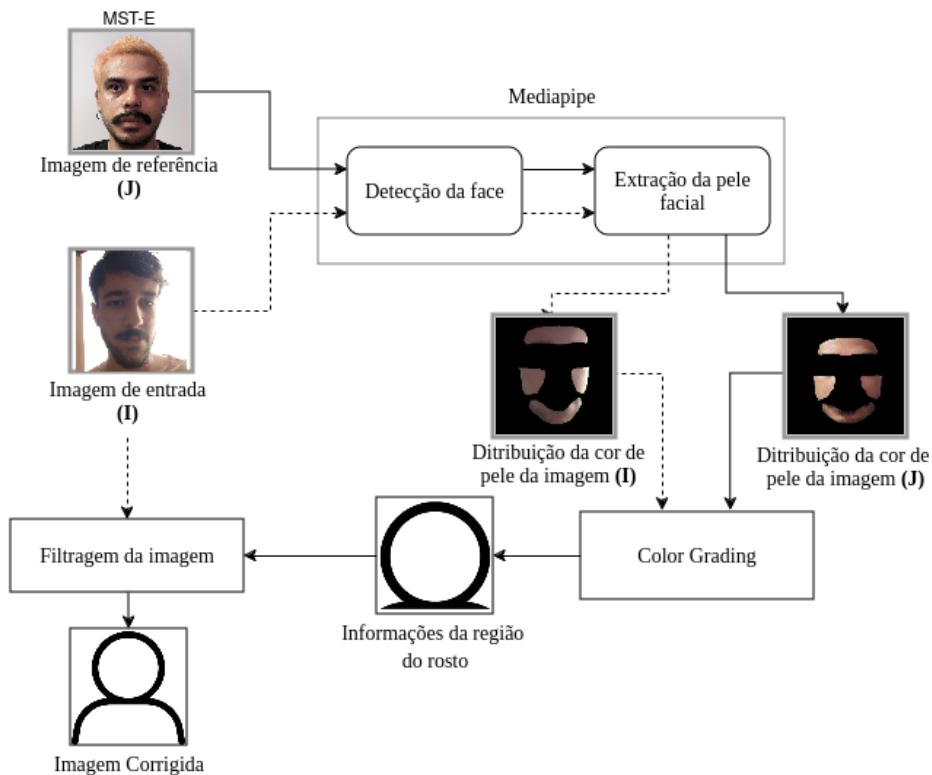
A proposta de correção de cor de uma imagem I emprega uma imagem de referência J . A imagem J é tida como idealmente precisa e possui uma cor representativa da escala de cor adotada. Este modelo de *whitebalance* é inspirado no trabalho de Shirai et al. (2021). A realização desta tarefa implica na execução das seguintes etapas:

- Detectar a face de pessoas em imagens;
- Extrair características referentes à pele facial de imagens;
- Realizar gradações de cor (Color Grading), promovendo uma transição suave do tom de pele da imagem I segundo o tom de pele da imagem de referência J ;
- Aplicar uma filtragem para redução de ruídos;
- Fazer a fusão do novo tom de pele da imagem I com os elementos de fundo existentes (que não sejam pele) na imagem de entrada.

Para realizá-las, as seguintes atividades foram desenvolvidas: a) estudo dos mecanismos do MediaPipe; b) estudo das técnicas de Color Grading; c) extração das características de pele facial; d) aplicação do Color Grading; e) filtragem das imagens obtidas e f) implementação final da correção de cor com a combinação dos algoritmos.

O diagrama da Figura 1 apresenta os principais módulos constituintes da metodologia proposta. O cronograma, como definido originalmente, está representado na Tabela 1 (atualizado conforme descrito no último relatório).

Figura 1 – Diagrama referente a metodologia utilizada para correção de cor de tons de pele.



Fonte: Elaborada pelo autor.

Atividades	Meses de trabalho					
	1º e 2º	3º e 4º	5º e 6º	7º e 8º	9º e 10º	11º e 12º
Estudo dos mecanismos do MediaPipe	•	•				
Estudo das técnicas de Color Grading		•	•			
Extração das características de pele facial		•	•			
Aplicação do Color Grading		•	•			
Redação do relatório parcial			•			
Filtragem das imagens obtidas			•	•		
Implementação final da correção de cor				•	•	
Avaliação e testes				•	•	•
Desenvolvimento do relatório final					•	•

Tabela 1 – Cronograma de atividades para 12 meses de trabalho.

O cronograma apresentado foi cumprido como originalmente proposto. As próximas seções deste relatório detalham a finalização do projeto, e descrevem como as atividades previstas na Tabela 1, no período decorrido, foram executadas. A seção 2 apresenta as principais atividades realizadas na segunda metade do projeto. A seção 3 expõe os resultados obtidos e as análises levantadas sobre eles. Por fim, a seção 4 apresenta as considerações levantadas para a conclusão do processo desta iniciação científica.

2 Atividades realizadas

Nesta seção estão relatadas as atividades desenvolvidas nesta segunda metade do projeto. A seção 2.1 detalha o processo de aplicação de uma filtragem bilateral para a redução dos ruídos comentados no último relatório. Já a seção 2.2 descreve como foi realizada a implementação final da correção de cor, com a combinação dos algoritmos desenvolvidos em cada atividade proposta.

2.1 Filtragem Bilateral

O filtro bilateral (PARIS et al., 2009) é um filtro de suavização não linear que preserva bordas, comumente utilizado para redução de ruídos em imagens. Este filtro substitui a intensidade de cada pixel por uma média ponderada das intensidades dos pixels vizinhos. Esse peso pode ser calculado com base em uma distribuição Gaussiana. Com isso, o filtro bilateral é definido como:

$$I_{\text{filtrado}}(x) = \frac{1}{W_p} \sum_{x_i \in \Omega} I(x_i) f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|),$$

e o termo de normalização, W_p , é definido como:

$$W_p = \sum_{x_i \in \Omega} f_r(\|I(x_i) - I(x)\|) g_s(\|x_i - x\|),$$

onde:

- I_{filtrado} representa a imagem após o filtro;
- I é a imagem original;
- x denota as coordenadas do pixel em processamento;
- f_r é a função de intensidade, utilizada para suavizar diferenças nos valores de intensidade (geralmente uma Gaussiana);
- g_s é a função espacial ou de domínio, que suaviza variações nas posições espaciais (também comumente uma Gaussiana);
- Ω é a janela centrada no pixel x , e $x_i \in \Omega$ refere-se aos demais pixels da janela.

Para remoção dos ruídos presentes nos resultados obtidos na primeira parte do projeto, foi escolhida a utilização de uma Filtragem Bilateral, visando a preservação das bordas na imagem corrigida. No entanto, como apresentado, o filtro bilateral não é isento de parâmetros, a saber:

- d : Diâmetro da vizinhança (tamanho da janela).
- σ_{color} : Parâmetro de suavização baseada em diferença de intensidade, presente na função f_r .
- σ_{space} : Parâmetro de suavização baseada em distância espacial, contido na função g_s .

Tendo em vista tal característica do filtro bilateral, foi elaborado um processo que permitisse a estimativa de tais parâmetros de forma “automática”. Este processo utilizou uma heurística baseada na razão entre o tamanho da área da face humana e a área total da imagem.

A área da face humana foi calculada como o número de pixels não nulos na imagem segmentada, denotada por $face_area$, enquanto a área total da imagem foi representada por $total_area$. Assim, a razão entre a área da face e a área total, chamada de $face_ratio$, foi definida como:

$$face_ratio = \frac{face_area}{total_area}$$

Com base nesse parâmetro, os valores dos parâmetros do filtro bilateral foram ajustados dinamicamente da seguinte forma:

- Tamanho da janela d : foi definido como o maior valor entre um limite mínimo de 5 e $15 \cdot face_ratio$, correspondentes ao tamanho da janela em pixels arredondado para o inteiro mais próximo:

$$d = \max(5, \lfloor 15 \cdot face_ratio \rfloor)$$

- Desvios padrão σ_{color} e σ_{space} : foram definidos como no mínimo 50 acrescentados de $50 \cdot face_ratio$, arredondados para o inteiro mais próximo:

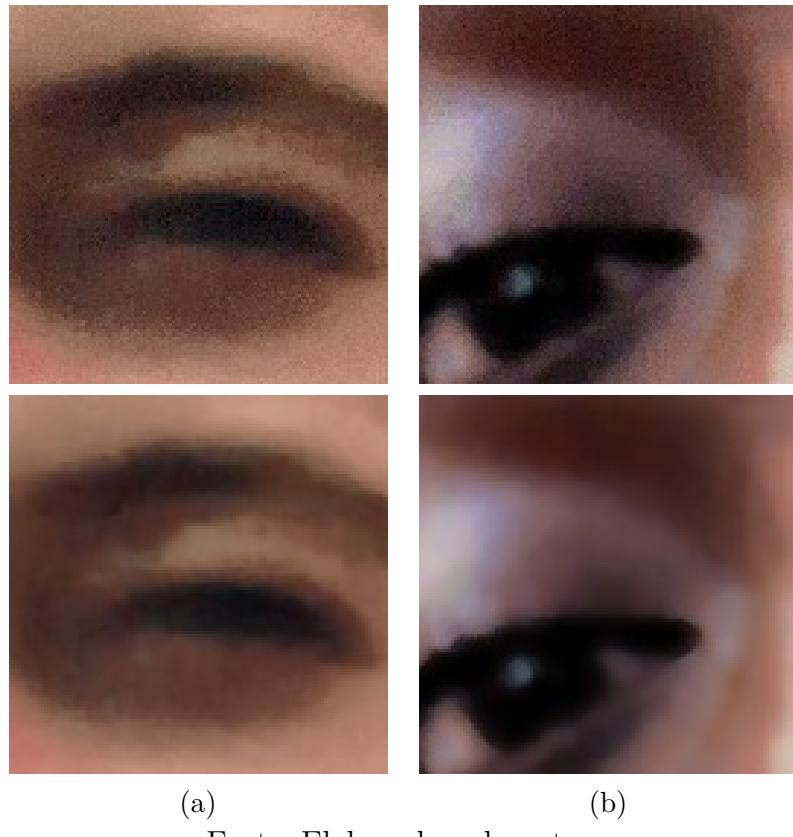
$$\sigma_{color} = \sigma_{space} = 50 + \lfloor 50 \cdot face_ratio \rfloor$$

A escolha desses ajustes dinâmicos foi baseada no fato de que imagens com áreas de face segmentadas menores requerem parâmetros menores para preservar detalhes importantes. Por outro lado, imagens com áreas de face maiores permitem o uso de parâmetros maiores, favorecendo uma suavização mais intensa em regiões maiores e resultando em melhores correções.

Essa heurística garante que o filtro bilateral se adapte de forma automática às características das imagens, eliminando a necessidade de ajuste manual dos parâmetros para cada imagem, permitindo ajustar o filtro bilateral de acordo com as características específicas da imagem.

Os resultados obtidos após a aplicação da Filtragem Bilateral podem ser analisados na Figura 2. É possível observar que artefatos gerados pela aplicação do Color Grading são reduzidos após a aplicação do filtro bilateral.

Figura 2 – Resultados da aplicação da Filtragem Bilateral, apresentando as imagens sem o filtro na linha superior e as imagens filtradas na linha inferior, para um tom de pele mais claro (a) e um mais escuro (b).



Fonte: Elaborada pelo autor.

2.2 Integração dos Algoritmos

Nesta etapa do projeto foi desenvolvido um *notebook* em Python que integra a implementação de todas as etapas do desenvolvimento apresentadas. A versão final (CASTRO, 2024) está disponível no GitHub¹.

Dentre as principais ferramentas e tecnologias empregadas, destaca-se a biblioteca OpenCV (2000), que foi essencial para a aplicação dos procedimentos relacionados às imagens, como o uso da Filtragem Bilateral, e a utilização do *framework* Mediapipe (LUGARESI et al., 2019), desenvolvido pelo Google, para a detecção facial e extração das características do rosto humano. O código com *notebook* está organizado conforme descrito nas seguintes seções:

¹ <https://github.com/rodrigorcz/skin-color-correction>

1. **Bibliotecas e Funções Auxiliares:** apresentação das tecnologias utilizadas;
2. **Detecção e Extração de Regiões de Interesse (ROI):** utilização do Mediapipe para detecção facial e extração de regiões de interesse da face humana;
3. **Color Grading:** aplicação das técnicas de Color Grading descritas por Pitié, Kokaram e Dahyot (2007);
4. **Filtro Bilateral:** aplicação de uma filtragem bilateral conforme descrito na seção anterior 2.1;
5. **Avaliação e Testes:** apresentação dos resultados e suas avaliações do método de correção de cor desenvolvido, que serão descritas na seção 3.

3 Resultados obtidos

Inicialmente é necessário destacar que todas as imagens de face humana utilizadas nesta seção são parte de um conjunto de exemplos fornecidos pelo Google (SCHUMANN et al., 2023), referentes à escala de cor de pele *Monk Skin Tone* (MST) (MONK, 2023), a qual foi utilizada neste projeto e que está representada na Figura 3.

Figura 3 – Escala MST para a classificação dos tons de pele



Fonte: Monk (2023)

A seguir, na Figura 4, é apresentado um exemplo de um resultado obtido após a aplicação da correção de cor proposta. Nesse exemplo, há uma imagem alvo que faz parte de um conjunto de *golden-images*, que são imagens que melhor caracterizam os tons de pele da escala *Monk Skin Tone*, ou seja, são tidas como imagens com tons de pele idealmente precisos. Essa imagem é utilizada para corrigir a cor de pele facial de uma imagem de entrada de uma pessoa que está, inicialmente, coberta por sombras. Após a aplicação do método proposto, é gerada uma imagem resultante, na qual há uma melhoria significativa na representação original da cor facial da pessoa.

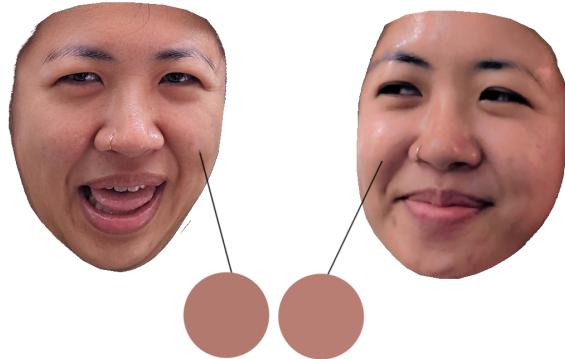
Figura 4 – Resultados obtidos para correção de cor de pele facial entre uma imagem de referência (a) e uma imagem de entrada (b), resultando em uma imagem (c) com a cor pele correta.



Fonte: Elaborado pelo Autor.

É importante ressaltar que a correção de cor na imagem é aplicada somente à região da face, o que confere à imagem resultante um aspecto “não natural” da cor da face com relação ao restante da imagem, notadamente influenciado pelo efeito de sombreamento da cena. No entanto, a cor de pele da imagem resultante possui uma grande similaridade com a cor da imagem alvo, como demonstrado na Figura 5.

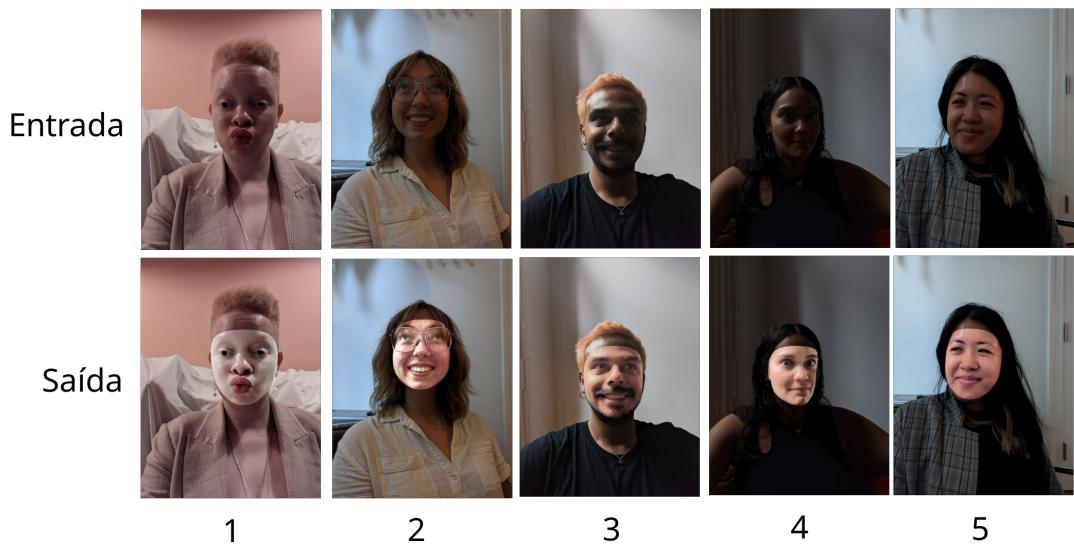
Figura 5 – Aproximação da cor da face segmentada referente a imagem alvo e a imagem corrigida.



Fonte: Elaborada pelo autor.

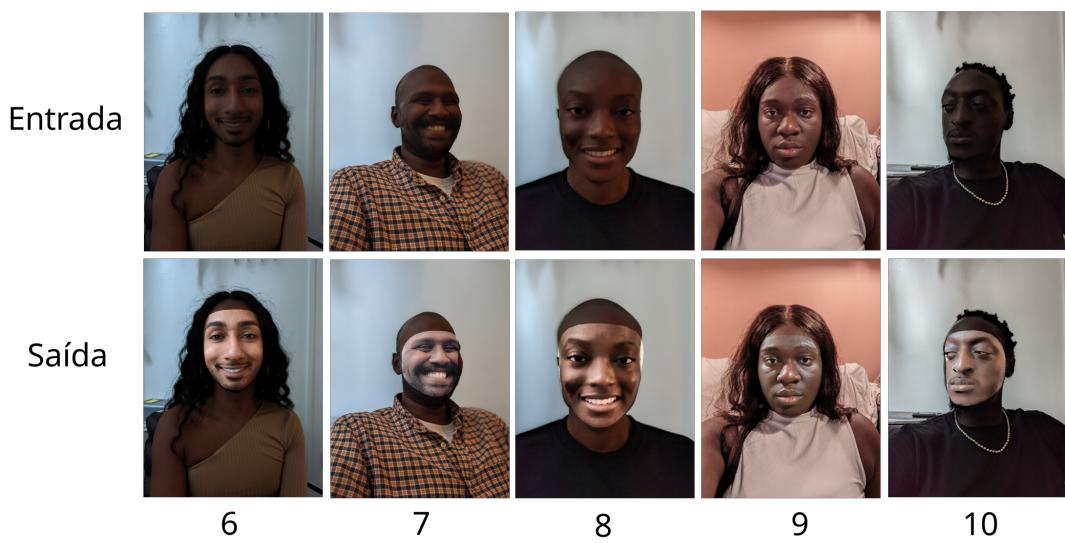
A seguir, as Figuras 6 e 7 apresentam os resultados obtidos para a correção de cor de pele em imagens de entrada cobertas por sombras, para indivíduos com todos os tons de pele na escala MST, tendo como imagem alvo as *golden-images* mencionadas anteriormente. Nessas imagens, é possível analisar um aperfeiçoamento considerável na reprodução fiel da cor facial original da pessoa.

Figura 6 – Resultados obtidos para correção cor de pele em imagens com indivíduos das escalas MST de 1 a 5.



Fonte: Elaborada pelo autor.

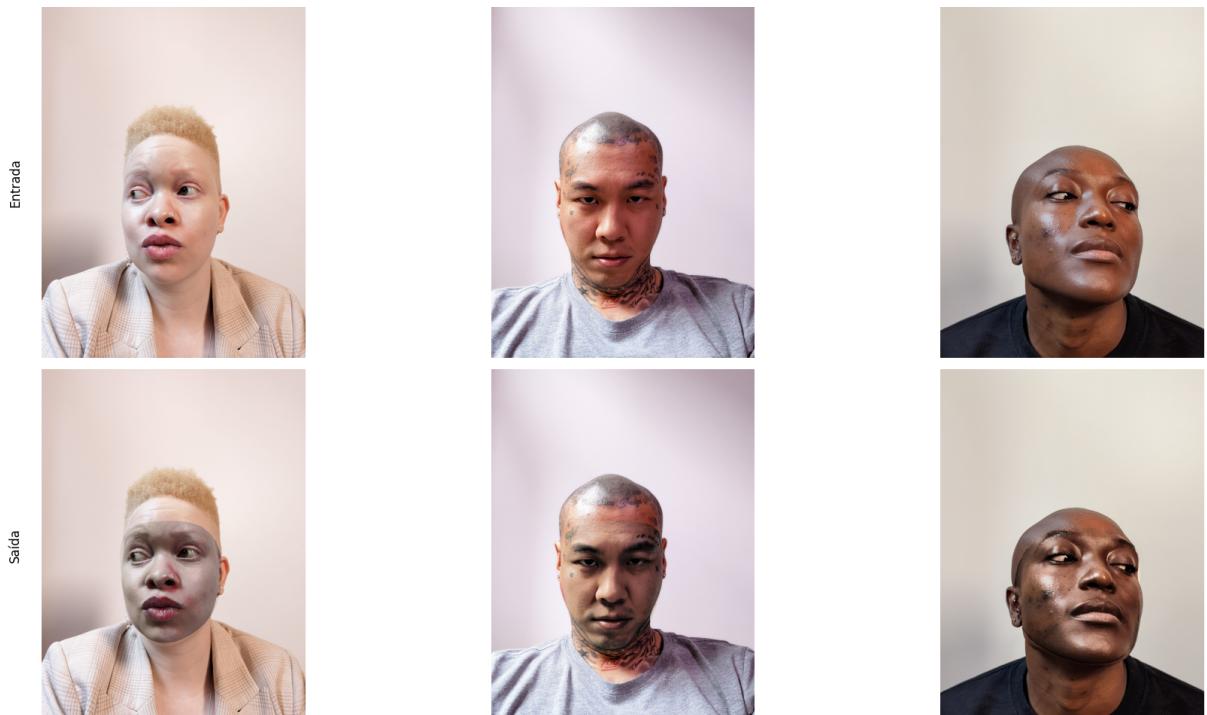
Figura 7 – Resultados obtidos para correção cor de pele em imagens com indivíduos das escalas MST de 6 a 10.



Fonte: Elaborada pelo autor.

Para casos em que as imagens de entrada estão influenciadas por fontes de luz externa, como flashes de celular ou raios solares, a metodologia também demonstrou eficiência, conforme ilustrado na Figura 8.

Figura 8 – Resultados obtidos para correção cor de pele em imagens com fontes de luz externa. A primeira linha apresenta as imagens de entrada e a segunda linha as imagens após a correção de cor.



Fonte: Elaborada pelo autor.

Devido a natureza extremamente subjetiva da interpretação de cor de pele pelo sistema visual humano, neste projeto não empregamos uma avaliação quantitativa dos resultados.

No entanto, qualitativamente, observamos que os resultados obtidos nos testes mostraram-se promissores para todos os tons de pele em imagens sob diversas influências em sua coloração, conforme apresentado nas Figuras 6, 7 e 8.

É importante mencionar, entretanto, que o método proposto possui algumas limitações que impedem a correta correção de cor em certos contextos. A Figura 9 apresenta uma situação inicial onde a presença de objetos na face, tais como óculos e máscaras, gerou correções imprecisas na imagem, sobrepondo a cor original dos objetos pela cor de pele da imagem alvo.

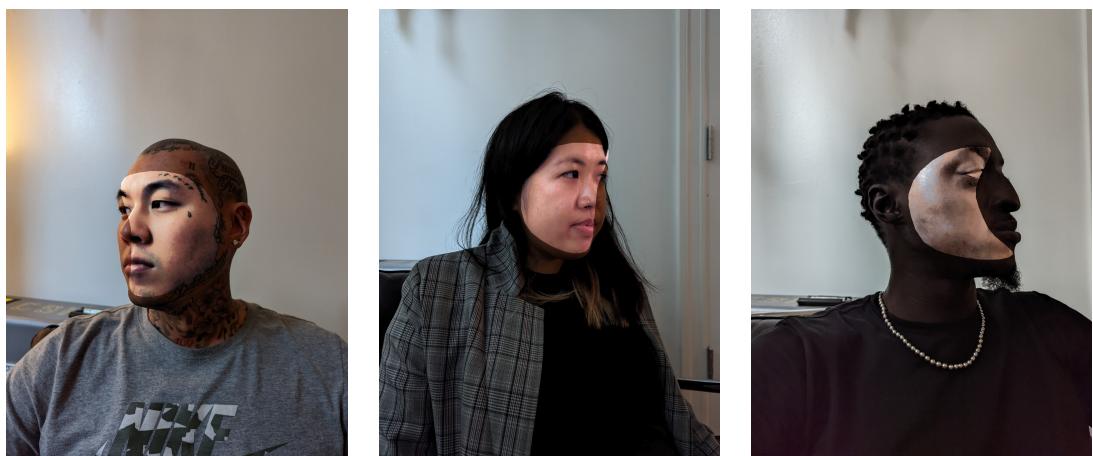
Figura 9 – Correção imprecisa de cor devido a presença de objetos na face.



Fonte: Elaborado pelo Autor.

Outro fator que limita o projeto está ligado à extração das características faciais. Quando o indivíduo apresenta inclinações na cabeça, a correção não é aplicada de maneira uniforme em todo o rosto. Essa limitação decorre das restrições do *framework* Mediapipe, utilizado para realizar a extração das características faciais. A Figura 10 ilustra algumas situações em que a correção de cor não foi adequadamente aplicada ao rosto humano devido a essa restrição.

Figura 10 – Resultados ruins gerados por limitações na metodologia de extração de características faciais.



Fonte: Elaborado pelo Autor.

4 Considerações Finais

Nesta seção, serão realizadas discussões finais referentes ao desenvolvimento geral do projeto e aos principais resultados obtidos ao longo do desenvolvimento deste trabalho, com uma análise crítica sobre suas implicações e limitações.

A metodologia desenvolvida neste trabalho baseou-se no uso de uma imagem de referência idealmente precisa para corrigir a coloração de uma imagem de entrada. Por meio da segmentação das regiões de interesse da pele em ambas as imagens, foram aplicadas técnicas de gradação de cor que garantiram uma transição suave entre diferentes tonalidades de pele, transferindo as cores da imagem precisa para a imprecisa. Esse processo permitiu corrigir a coloração da pele na imagem afetada, preservando a fidelidade das cores no restante da composição.

Os resultados apresentados na Seção 3 evidenciaram que a abordagem proposta é eficaz para corrigir tons de pele em imagens submetidas a diferentes condições de coloração. Esses resultados reforçam que os objetivos iniciais foram alcançados, com esses dados que comprovam a eficiência da metodologia proposta.

Embora os resultados tenham alcançado boa concordância com as expectativas iniciais, alguns desafios foram encontrados, tais como a presença de objetos na face e as limitações envolvidas com a etapa de extração das características faciais, os quais impactaram negativamente a eficiência da correção.

Assim, os resultados apresentados têm o potencial de impactar positivamente áreas de pesquisa onde imagens faciais são os meios de análise, tais como, fotografia, dermatologia, cosméticos e entretenimento visual, por exemplo, possibilitando que mesmo com vários fatores externos, como irregularidades naturais da pele, divergências nas fontes de iluminação e variação na configuração do dispositivo de captura, as imagens possuam as cores de pele proporcionais às reais.

Por fim, é válido mencionar a participação deste projeto no Simpósio Internacional de Iniciação Científica e Tecnológica da USP (SIICUSP) 2024, realizado em outubro. Durante o evento, o bolsista teve a oportunidade de apresentar os resultados da pesquisa por meio de um pôster e apresentações orais, contribuindo para a divulgação científica.

Referências

- CASTRO, R. R. *Skin Color Correction*. GitHub, 2024. Disponível em: <<https://github.com/rodrigorcz/skin-color-correction>>.
- LUGARESI, C. et al. Mediapipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*, 2019.
- MONK, E. The monk skin tone scale. SocArXiv, 2023.
- OPENCV. 2000. Disponível em: <<https://opencv.org>>. Acesso em: 20 feb 2021.
- PARIS, S. et al. Bilateral filtering: Theory and applications. *Foundations and Trends® in Computer Graphics and Vision*, Now Publishers, Inc., v. 4, n. 1, p. 1–73, 2009.
- PITIÉ, F.; KOKARAM, A. C.; DAHYOT, R. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding*, v. 107, n. 1, p. 123–137, 2007. ISSN 1077-3142. Special issue on color image processing.
- SCHUMANN, C. et al. Consensus and subjectivity of skin tone annotation for ml fairness. *arXiv preprint arXiv:2305.09073*, 2023.
- SHIRAI, K. et al. Guided facial skin color correction. *Signals*, MDPI, v. 2, n. 3, p. 540–558, 2021. Disponível em: <<https://doi.org/10.3390/signals2030033>>.