

# Pesquisa e Desenvolvimento de Métodos de Visão Computacional para Mensuração de Movimentos Faciais

Murilo de Souza Freitas \*  
[murilo.sf1@puccampinas.edu.br](mailto:murilo.sf1@puccampinas.edu.br)  
Pontifícia Universidade Católica de Campinas  
Campinas, São Paulo, Brasil

Ademar Takeo Akabane  
[ademar.akabane@puc-campinas.edu.br](mailto:ademar.akabane@puc-campinas.edu.br)  
Pontifícia Universidade Católica de Campinas  
Campinas, São Paulo, Brasil

Alexandre Fonseca Brandão  
[alexandre.brandao@puc-campinas.edu.br](mailto:alexandre.brandao@puc-campinas.edu.br)  
Pontifícia Universidade Católica de Campinas  
Campinas, São Paulo, Brasil

Wemerson Delcio Parreira  
[wemerson.delcio@puc-campinas.edu.br](mailto:wemerson.delcio@puc-campinas.edu.br)  
Pontifícia Universidade Católica de Campinas  
Campinas, São Paulo, Brasil

## Abstract

This work presents the development of a computer vision solution for the quantitative assessment of facial movements, with the aim of supporting the diagnosis and rehabilitation of facial paralysis. The system, implemented in Python using the OpenCV and MediaPipe libraries, detects 468 facial landmarks and performs geometric calculations to monitor the amplitude and symmetry of expressions in real-time, using only a conventional webcam without depth sensors. Preliminary results demonstrate the tool's capability to translate vector coordinates into clinical metrics and instant visual feedback, establishing a foundation for future gamification applications in virtual environments aimed at stimulating neuroplasticity and patient engagement in physical therapy treatment.

## Keywords

Mensuração facial, Visão Computacional, Monitoramento, MediaPipe.

## 1 Introdução

A expressão facial constitui um dos principais meios de comunicação não verbal e desempenha papel fundamental tanto nas interações humanas quanto em contextos clínicos, especialmente na avaliação de distúrbios neuromusculares, como a paralisia facial [1]. A mensuração precisa desses movimentos é essencial para apoiar o diagnóstico, monitorar a evolução clínica e avaliar a eficácia de tratamentos de reabilitação [2], bem como para a tecnologia de sistemas de interação humano-computador. Contudo, métodos tradicionais de análise facial frequentemente dependem de equipamentos especializados, alto custo, o que limita sua ampla adoção em ambientes clínicos e domiciliares. [3].

Nesse contexto, técnicas de visão computacional têm se mostrado alternativas promissoras para a análise automatizada de movimentos faciais, permitindo a extração de informações cinemáticas a partir de imagens e vídeos. Avanços recentes em modelos de detecção de *landmarks* faciais (pontos-chave) permitem o rastreamento preciso de pontos anatômicos em tempo real. Entre essas soluções, a biblioteca MediaPipe destaca-se por oferecer modelos eficientes, de alta precisão e baixo custo computacional [4].

Este trabalho investiga o uso de *landmarks* faciais para a mensuração quantitativa de movimentos, como a abertura da boca, as piscadas e a amplitude do sorriso. Para quantificar a movimentação, utiliza-se o cálculo da diferença absoluta ou da distância euclidiana entre dois pontos de interesse na malha facial. Por exemplo, a lógica matemática fundamental para a detecção de abertura (boca ou olhos) é dada pela equação 1:

$$\text{Gap} = |y_{\text{inferior}} - y_{\text{superior}}| \quad (1)$$

Em que  $y$  representa a coordenada vertical dos pontos anatômicos rastreados. A partir desse valor, é possível determinar estados lógicos (como "Boca Aberta" ou "Piscando") com base em limiares pré-definidos, contribuindo para a quantificação da recuperação motora.

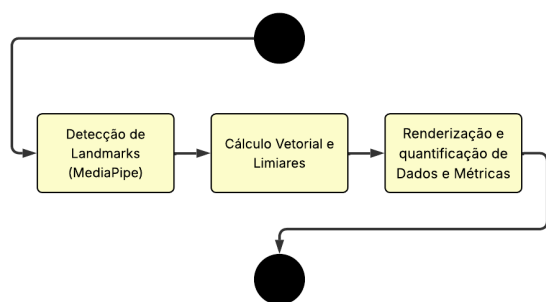
Além da quantificação métrica, a vetorização dos movimentos faciais abre caminho para aplicações na Realidade Estendida. A transformação de expressões faciais em coordenadas digitais permite que esses dados sejam enviados a motores gráficos, como o Unity, e que pacientes controlem objetos virtuais por meio de movimentos faciais. Essa abordagem visa estimular a neuroplasticidade e aumentar a adesão ao tratamento por

meio da gamificação, tornando a fisioterapia menos monótona e mais interativa. Dessa forma, este trabalho contribui para o desenvolvimento de um protótipo de baixo custo e de arquitetura modular, voltado ao apoio clínico e à integração futura com plataformas de reabilitação digital, reforçando o potencial da visão computacional como ferramenta na informática em saúde.

## 2 Metodologia

Este trabalho adota uma abordagem computacional para a mensuração quantitativa de movimentos faciais baseada em videogrametria, utilizando técnicas de visão computacional em tempo real. O sistema foi projetado com foco em acessibilidade e baixo custo, operando em computadores pessoais convencionais e utilizando exclusivamente uma *webcam*, sem a necessidade de sensores de profundidade ou hardware especializado.

A solução foi implementada em Python, integrando a biblioteca OpenCV para a captura e o pré-processamento dos quadros de vídeo e o framework MediaPipe Face Mesh para a detecção de *landmarks* faciais. O MediaPipe foi selecionado por oferecer modelos pré-treinados robustos, capazes de inferir 468 pontos anatômicos tridimensionais ( $x, y, z$ ) da face com baixa latência e alta precisão, características essenciais para aplicações interativas e clínicas em tempo real. O fluxo de processamento de dados, desde a captura da imagem até a exibição das métricas, é apresentado esquematicamente na **Figura 1**.



**Figura 1: Diagrama de atividades que representa o fluxo de processamento do algoritmo: detecção da malha, cálculo vetorial e renderização das métricas.**

Conforme ilustrado na Figura 1, o algoritmo opera em um ciclo contínuo (*loop*) dividido em três etapas principais:

**Detecção de Landmarks:** O sistema recebe o quadro de vídeo (frame), converte o espaço de cores (BGR para RGB) e aplica o modelo de aprendizado de

máquina do MediaPipe para inferir 468 *landmarks* tridimensionais na face do usuário;

**Cálculo Vetorial e Limiares:** Com as coordenadas normalizadas ( $x, y, z$ ) dos pontos de interesse (lábios, pálpebras, íris e contorno facial), o algoritmo realiza cálculos de geometria analítica (distâncias euclidianas e diferenças absolutas) para quantificar a amplitude dos movimentos. Limiares pré-definidos são aplicados para classificar estados lógicos, como "Boca Aberta", "Sorrindo" e "Piscando".

**Renderização e Quantificação:** Por fim, os dados quantitativos e o *feedback* visual (malha sobreposta e marcadores coloridos) são renderizados na interface gráfica em tempo real, permitindo o monitoramento instantâneo pelo profissional de saúde ou pelo paciente.

### 2.1 Implementação

Para a mensuração dos movimentos faciais, foram selecionados *landmarks* específicos associados a regiões anatômicas clinicamente relevantes. A escolha desses pontos baseou-se em dois critérios principais: (i) estabilidade do rastreamento ao longo do tempo e (ii) relação direta com a mobilidade facial voluntária e involuntária. Por exemplo, os *landmarks* 13 (lábio superior) e 14 (lábio inferior) foram utilizados para a mensuração da abertura da boca, pois representam pontos centrais e simétricos da região labial, minimizando os efeitos de rotação da cabeça e de variações laterais. A implementação da captura e conversão de coordenadas para a mensuração é detalhada no trecho de código a seguir, onde as coordenadas normalizadas do MediaPipe são convertidas para pixels da imagem ( $w, h$ ):

```
# Cálculo simplificado da abertura vertical
cima_y = int(landmarks[13].y * h)
baixo_y = int(landmarks[14].y * h)
boca_gap = abs(baixo_y - cima_y)

# Classificação de estado
if boca_gap > 20: # Limiar em pixels
    status = "Boca Aberta"
```

De forma análoga, os cantos esquerdos (61) e direitos (291) da boca foram selecionados para avaliar a amplitude do sorriso, enquanto *landmarks* nas pálpebras superiores e inferiores foram empregados na detecção de piscadas.

A classificação dos estados funcionais é realizada por meio da comparação dessas métricas com limiares empíricos previamente definidos. Por exemplo, o parâmetro BOCA\_ABERTA\_LIMITE = 20 (em *pixels*) foi determinado a partir de testes exploratórios em

ambiente controlado, considerando diferentes usuários e condições de iluminação, a fim de identificar um valor mínimo capaz de distinguir consistentemente os estados de boca aberta e fechada. Procedimento semelhante foi adotado para os limiares associados à detecção de piscadas e de sorrisos. Embora eficazes nos testes preliminares, esses limiares dependem da escala da imagem, o que motiva, como trabalho futuro, a implementação de estratégias de normalização e de calibração métrica.

### 3 Resultados Preliminares

O protótipo desenvolvido foi avaliado por meio de testes preliminares em ambiente controlado de laboratório, utilizando uma *webcam* convencional para o rastreamento facial em tempo real. Durante a execução, o sistema demonstrou capacidade estável de detectar a face do usuário e inferir corretamente os *landmarks* faciais, mantendo rastreamento contínuo mesmo sob variações moderadas de iluminação e sob pequenos movimentos da cabeça.

Diferentemente de abordagens estáticas, o sistema fornece *feedback* contínuo na tela, exibindo métricas calculadas instantaneamente. Conforme demonstrado nos testes de execução, o sistema exibe métricas sobrepostas à imagem, incluindo a Direção da Cabeça (baseada na posição do nariz), o status da Boca (Sorrindo, Bico, Aberta), o monitoramento de Olhos (abertos/fechados com contagem de piscadas) e a inferência de Reação emocional básica (Feliz/Surpresa). Um exemplo de saída de dados do protótipo é apresentado na Figura 2.



**Figura 2: Janela de execução do programa, exibindo o rastreamento do rosto e a exibição das informações (Direção H: Centro; Direção V: Baixo; Boca: Sorrindo; Olhos: Abertos; Número de piscadas: 60)**

Os resultados obtidos, embora preliminares e de natureza exploratória, confirmam a viabilidade da abordagem proposta para converter as coordenadas geométricas de *landmarks* faciais em métricas

quantitativas relevantes. Esses achados reforçam o potencial do sistema como ferramenta de apoio à mensuração de movimentos faciais e fundamentam as próximas etapas de validação quantitativa e calibração métrica do método.

### 3.1 Limitações

As limitações atuais do projeto residem na dependência de limiares fixos (ex.: BOCA\_ABERTA\_LIMITE = 20) definidos em *pixels*. Isso significa que a precisão da classificação pode variar conforme a distância do usuário em relação à câmera (escala). Para uma aplicação clínica universal, será necessário implementar uma etapa de calibração que converta essas medidas em uma escala métrica relativa ao tamanho da face do usuário, mitigando o erro de perspectiva.

### 4 Considerações Finais

A identificação e mensuração de movimentos faciais por meio de visão computacional constituem o alicerce deste trabalho, demonstrando a viabilidade de ferramentas de baixo custo para auxiliar no diagnóstico. A integração de dados brutos para a reconstrução de estados lógicos evidencia o potencial da ferramenta. Em etapas posteriores, será implementada a normalização das distâncias com base na largura da face para mitigar erros de perspectiva.

Entretanto, o impacto desta pesquisa estende-se além da mera mensuração. Como próximas etapas, prevê-se a calibração métrica do sistema (convertendo pixels em milímetros) e a integração desses inputs com ambientes virtuais no Unity. O objetivo final é estabelecer um sistema de gamificação que estimule a musculatura facial de pacientes com paralisia, unindo a precisão dos dados quantitativos, essenciais ao acompanhamento clínico, ao engajamento lúdico necessário para uma reabilitação eficaz.

### Referências

- [1] Khabaralak, Kostiantyn & Koriashkina, Larysa. (2022). Fast Facial Landmark Detection and Applications: A Survey.
- [2] Dukić, David & Sovic Krzic, Ana. (2022). Real-Time Facial Expression Recognition Using Deep Learning with Application in the Active Classroom Environment. *Electronics*. 11. 1240.
- [3] Bandini A, et al. A New Dataset for Facial Motion Analysis in Individuals With Neurological Disorders. *IEEE J Biomed Health Inform*. 2021.
- [4] Barbosa J, et al. Efficient quantitative assessment of facial paralysis using iris segmentation and active contour-based key points detection with hybrid classifier. *BMC Med Imaging*. 2016.