



Classificação de Formato de Rostos

Alunos: João Campelo, Laura Calleja e Guilherme Duran

Disciplina: Visão Computacional

Instituição: Universidade IBMEC

Data: 2025

Introdução ao Projeto

Contexto e Motivação

A análise automatizada de características faciais representa um campo em expansão na visão computacional, com aplicações que vão desde sistemas de segurança até recomendações personalizadas em aplicativos de beleza e saúde. A classificação de formato de rostos é uma tarefa desafiadora que combina técnicas de detecção facial, extração de características geométricas e aprendizado de máquina.

Este projeto explora métodos computacionais para identificar e classificar rostos humanos em quatro categorias principais: **quadrado, oval, redondo e triangular invertido.**

Objetivo do Projeto

Desenvolver um sistema capaz de:

- Detectar rostos em imagens digitais com alta precisão
- Identificar landmarks faciais (pontos de referência) de forma robusta
- Classificar automaticamente o formato facial em uma das quatro categorias estabelecidas

Metodologia de Desenvolvimento

Pipeline do Sistema

O projeto foi estruturado em três etapas principais, cada uma desempenhando um papel fundamental no processo de classificação. A arquitetura do sistema foi projetada para ser modular, permitindo otimizações independentes em cada fase do processamento.



Detectção de Rosto

Identificação e localização do rosto na imagem utilizando algoritmos de detecção facial. Esta etapa isola a região de interesse e prepara a imagem para análise posterior.

Extração de Landmarks

Identificação de pontos faciais chave (landmarks) como contorno do rosto, posição do queixo, posição das maçãs do rosto e testa.

Classificação

Análise das características geométricas extraídas e classificação do formato facial em uma das quatro categorias: quadrado, oval, redondo ou "coração".

Considerações Iniciais

- Condições Ideais para Captura das Imagens
- Iluminação adequada e uniforme sobre o rosto;
- Limites faciais claramente visíveis (testa, mandíbula e queixo bem definidos);
- Pessoa posicionada a uma distância de até 2 metros da câmera;
- Postura neutra, com mínima rotação ou inclinação da cabeça.

Reconhecimento Facial

Fundamentos da Detecção

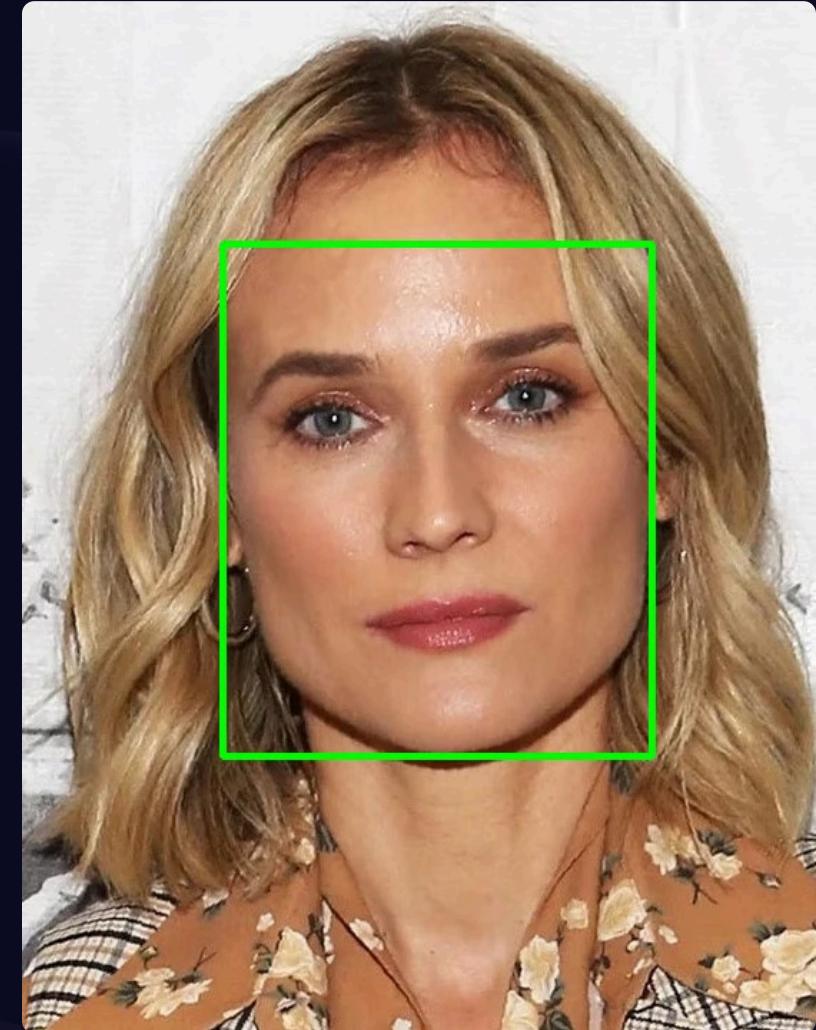
A detecção facial é o primeiro passo crítico no pipeline de classificação.

Extração de Landmarks Faciais:

Após a detecção do rosto, o sistema identifica os pontos de referência (landmarks) que mapeiam características faciais importantes. Estes pontos incluem:

- **Contorno facial**
- **testa**
- **zigoma**
- **queixo**

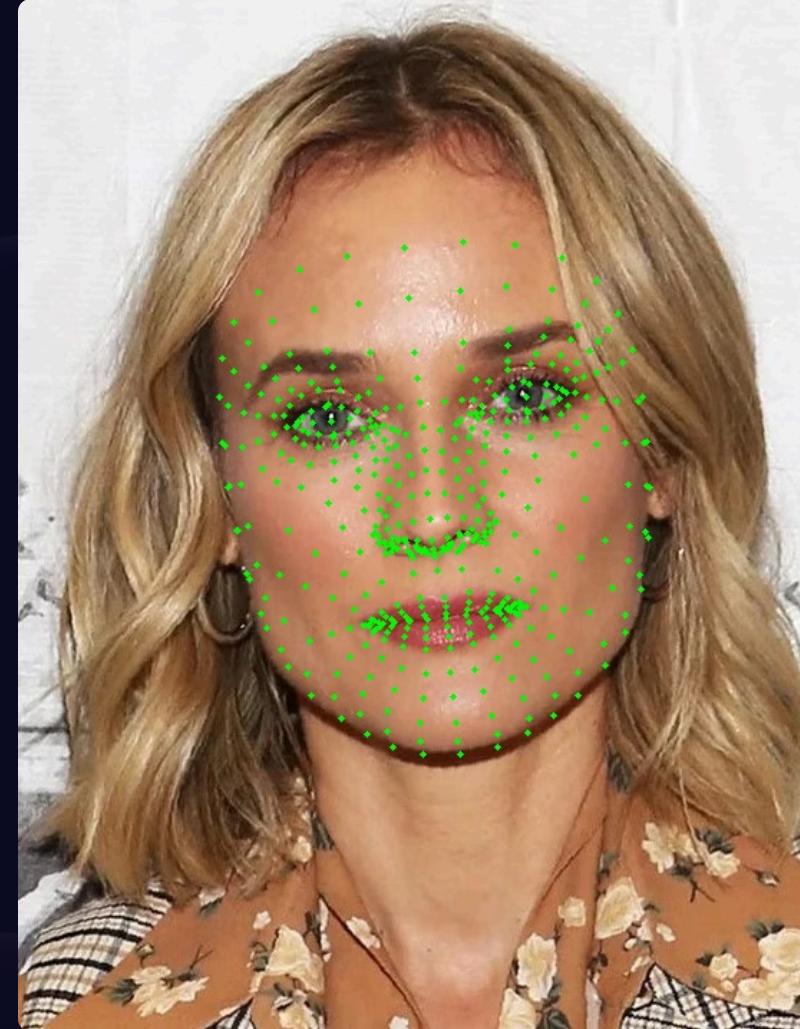
Estes landmarks servem como base para calcular medidas geométricas fundamentais, como razões de proporção facial, ângulos e distâncias relativas entre características.



Face Mesh: Detecção Detalhada de Landmarks Faciais

O algoritmo MediaPipe Face Mesh é uma peça central para a precisão do nosso sistema, permitindo a detecção de 468 pontos de referência faciais em 3D. Essa riqueza de detalhes possibilita uma análise geométrica muito mais refinada das proporções do rosto.

A capacidade de capturar uma rede tão densa de pontos, incluindo contornos da face, olhos, nariz e boca, é essencial para identificar as características distintivas de cada formato facial.



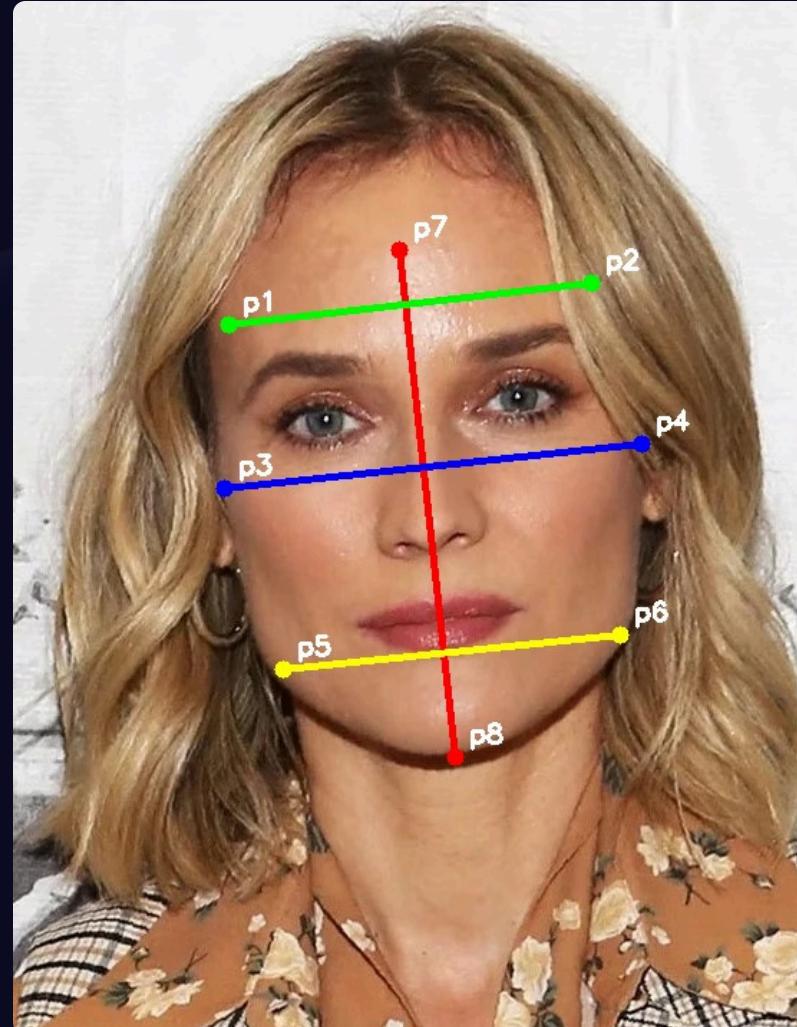
Cálculo de Features Geométricas

Para classificar os formatos faciais de forma precisa, selecionamos 8 landmarks faciais principais. A partir desses pontos estratégicos, calculamos distâncias e razões que são cruciais para a análise geométrica.

Distâncias Essenciais

São calculadas 4 distâncias fundamentais usando a **distância euclidiana** entre pares de landmarks, garantindo a medição precisa da geometria facial:

- **Largura da testa (Forehead width):** Mede a distância entre os pontos mais largos da testa.
- **Largura zigomática (Cheekbone width):** Corresponde à distância entre as maçãs do rosto.
- **Largura da mandíbula (Jawline width):** Mede a largura da parte mais proeminente da mandíbula.
- **Altura facial (Facial height):** A distância do topo da testa ao ponto mais baixo do queixo.



A partir dessas distâncias, derivamos 5 razões geométricas que são **normalizadas** e **invariantes à escala**, o que permite uma comparação robusta entre diferentes rostos, independentemente do tamanho ou da pose na imagem.

1

Razões Geométricas

R1 = Altura / Zigomático

R2 = Mandíbula / Testa

R3 = Mandíbula / Zigomático

R4 = Testa / Zigomático

R5 = Testa / Mandíbula

Dataset Utilizado

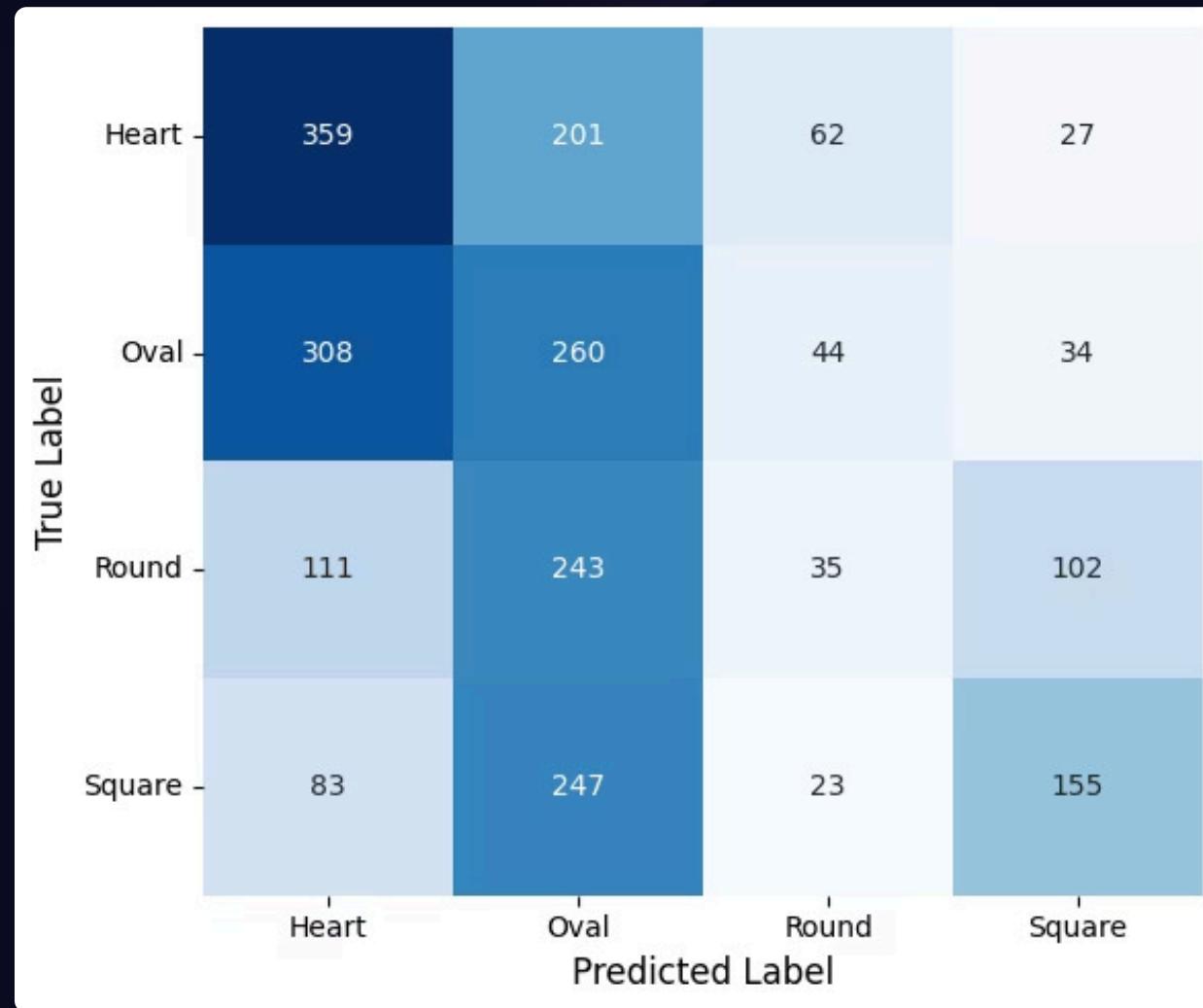
O conjunto de dados utilizado foi derivado do **Face Shape Dataset**, composto por cerca de 5.000 imagens frontais de celebridades femininas. Para este trabalho, selecionamos quatro categorias de formato facial, com **1.000 imagens por classe**, cada uma rotulada de acordo com sua morfologia.

Classificação de Formato de Rosto Utilizando Limiares

Face shape	$R_1 = \frac{d_h}{d_c}$	$R_2 = \frac{d_j}{d_f}$	$R_3 = \frac{d_j}{d_c}$	$R_4 = \frac{d_f}{d_c}$	$R_5 = \frac{d_f}{d_j}$
Heart	> 0.95	< 0.90	< 0.85	-	> 1.10
Oval	0.90–1.10	0.85–0.95	0.75–0.85	-	1.05–1.12
Round	≈ 1.00	-	-	≈ 0.90	1.10–1.20
Square	< 0.90	0.90–0.97	0.80–0.85	-	< 1.10

Resultado com Abordagem "Manual"

A acurácia utilizando esta abordagem foi de 20,25%, não foi capaz de identificar diferenças entre todas as classes



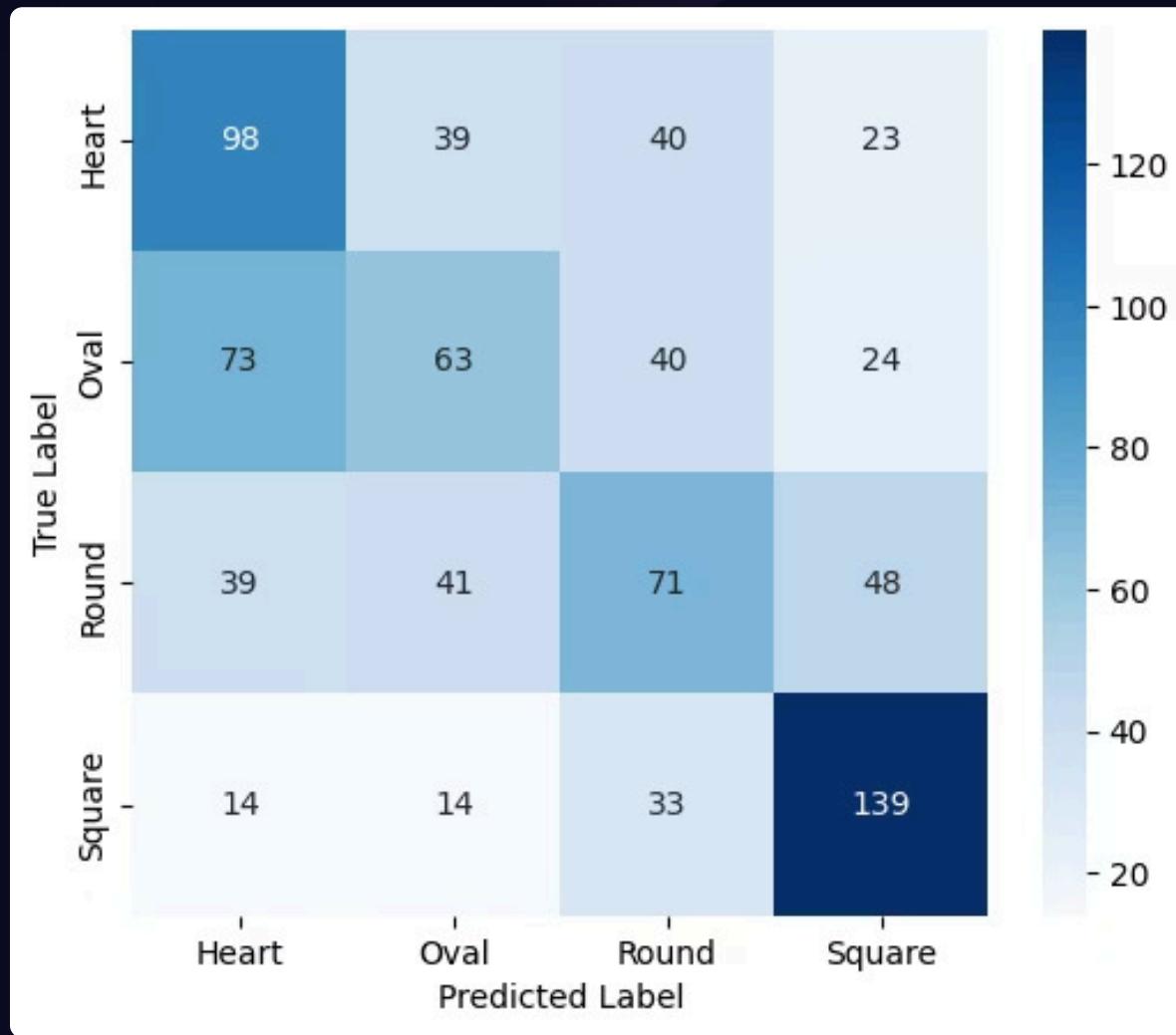
Classificação de Formato de Rosto com Machine Learning

Após calcular os cinco rácios geométricos extraídos dos pontos faciais, cada imagem passou a ser representada apenas por essas medidas numéricas. Com esse conjunto de características, foram treinados dois modelos de aprendizado supervisionado: **Random Forest** e **Gradient Boosting**. Esses algoritmos foram escolhidos por lidarem bem com relações não lineares e por apresentarem ótimo desempenho em dados tabulares de baixa dimensionalidade.

Os modelos foram comparados diretamente com o método baseado em limites geométricos apresentado anteriormente, permitindo avaliar a diferença entre uma abordagem **rule-based** e uma abordagem **data-driven**.

Resultado Random Forest

A acurácia utilizando esta abordagem foi de 50,00%, é possível notar uma melhor classificação entre as classes



Resultado Gradient Boosting

A acurácia utilizando esta abordagem foi de 48,25%, é possível notar uma melhor classificação entre as classes

