

# Detecção e Identificação de Caras com e sem Máscara

Elias Ferreira (up201502818), Nuno Minhoto (up201604509), Nuno Neto (up201506075), Lúcio Almeida (up201609022), Henrique Marques (up201605307)

**Resumo** – Neste documento é descrito o algoritmo desenvolvido e as metodologias adotadas para detectar e identificar caras e máscaras pertencentes a uma base de dados de imagens. A primeira parte envolve a detecção de caras numa imagem rgb com uma ou mais pessoas, e a criação de uma região de interesse. A segunda fase corresponde à detecção de máscaras faciais numa imagem já com a região da cara identificada, procedendo depois à avaliação de se esta está a ser utilizada corretamente ou não.

## I. DETECÇÃO DAS CARAS

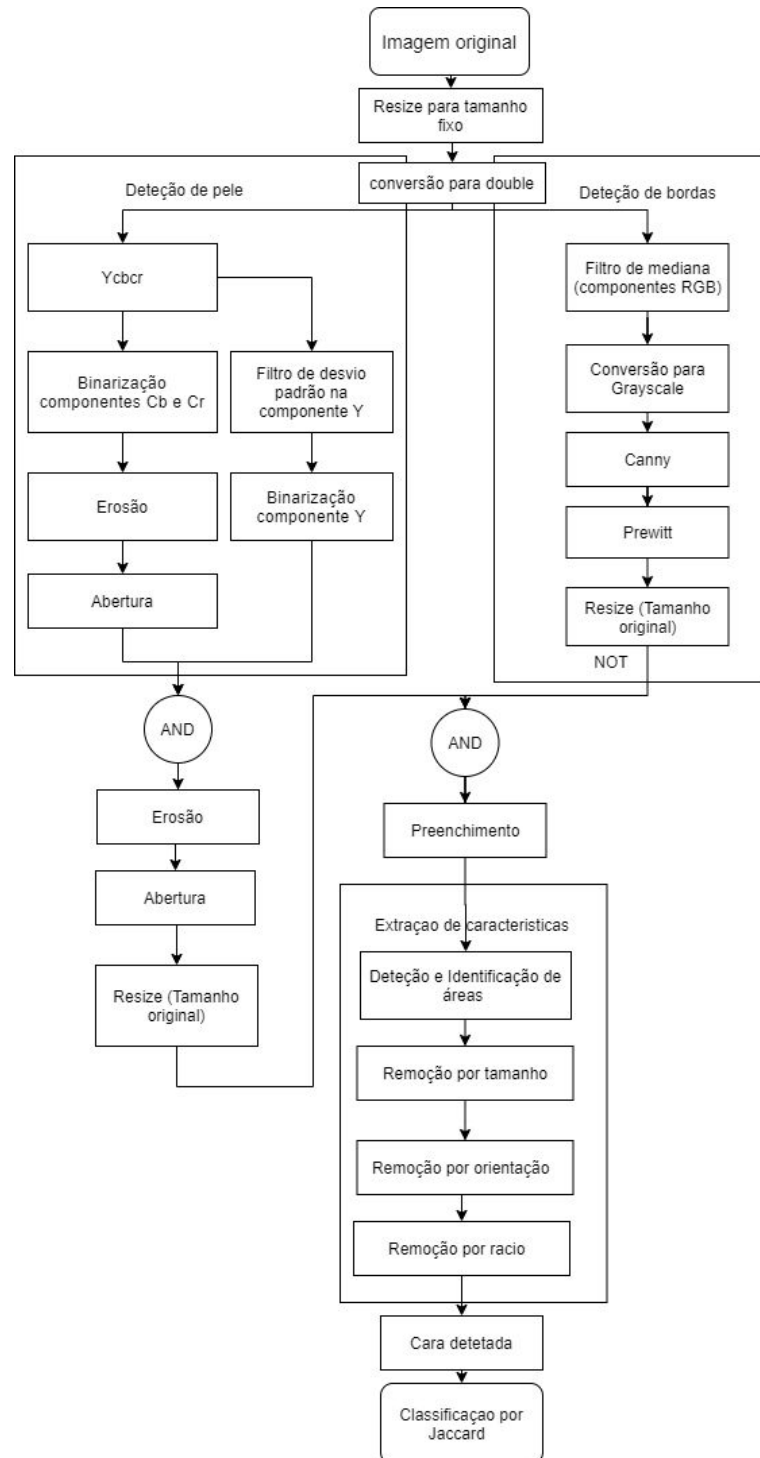
### A. Metodologia

Neste ponto vai ser feita a descrição e justificação do algoritmo desenvolvido para fazer a detecção e identificação das caras em cada imagem.

Em primeiro lugar foi feito um resize da imagem original para um tamanho fixo (1024 1559 que foi escolhido por ser o tamanho da maior imagem), de maneira a que todas as imagens analisadas tivessem o mesmo tamanho. Esta normalização melhora a performance geral, devido ao facto de as imagens variarem de tamanho e os elementos de erosão e abertura terem tamanho fixo. De seguida é feita uma conversão da imagem para double com o objectivo de aumentar a precisão das operações aritméticas efetuadas.

Após esta conversão são aplicados dois métodos em paralelo, um para realizar a detecção da pele nas imagens, e outro para a detecção das bordas das caras.

Na parte da detecção de pele, começamos por converter o espaço de cores original RGB para YCBCR. Devido ao facto de este standard variar linearmente as suas componentes, permitindo obter resultados mais previsíveis e fiáveis nas próximas operações. De seguida a imagem foi dividida nas suas componentes, fazendo posteriormente uma binarização das componentes que contém a informação de cor (Cb e Cr), com um threshold (Cr



entre 133/255 e 173/255, Cb entre 77/255 e 127/255).

De seguida as mesmas são sujeitas a uma erosão, com um elemento estruturante em cruz com tamanho 43, e uma abertura, com um disco de raio 20, de forma a realçar a informação realmente útil, que neste caso era regiões onde é mais provável conter pele. Por outro lado a componente de luminância (Y) foi submetida a uma filtragem de desvio padrão, seguido da sua binarização com threshold ( $Y \geq 1/255$ ), o que nos permitiu identificar os pontos com maior variação de luminância, que em condições ideais iriam corresponder a uma cara. De seguida foi feita a junção dos resultados destes dois métodos através de um AND, obtendo assim uma imagem binarizada com as regiões onde estariam as possíveis caras. As dimensões e tipos de elementos estruturantes, usados nas operações morfológicas anteriores foram obtidos através de várias tentativas e observação dos resultados, onde acabamos por escolher os valores aos quais correspondia um melhor resultado final. Os thresholds foram escolhidos baseados na bibliografia lida que foi fornecida pelos docentes. No fim destas operações foi feito um novo resize e a imagem foi devolvida a proporção original.



Figura 2. Resultado do método de identificação de pele

No outro método de detecção de bordas, primeiro começamos por aplicar um filtro de mediana às componentes RGB com o objetivo de reduzir as reflexões de luminosidade na cara, neste caso as mesmas podiam corromper as bordas que estamos interessados. Depois disso, procedemos à conversão para uma imagem em grayscale em preparação para aplicação do método de canny para a detecção de bordas, seguido de um operador de prewitt para reforçar as mesmas. Obtendo assim as bordas das caras, estes métodos já devolvem uma imagem binarizada.

No fim destas operações também foi feito um novo resize e a imagem foi devolvida a proporção original.



Figura 3. Resultado do método de identificação das bordas

A fusão destas duas componentes foi feita através da operação lógica AND da componente da cor de pele com a componente inversa das bordas, seguido de um preenchimento das regiões vazias.

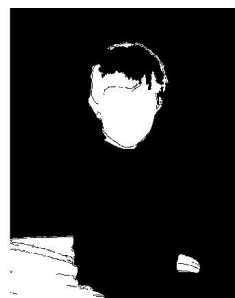


Figura 4. Resultado da junção

Para a extração de características, primeiro procedeu a identificação e detecção do número total de regiões e as suas características individuais, presentes na imagem binarizada final, usando as funções do Matlab.

Com isto, estas regiões passaram por uma seleção de requisitos que as elimina caso a probabilidade de ser uma cara seja baixa.

Primeiro todas as regiões consideradas demasiado pequenas (20 vezes mais pequenas que a maior região) foram eliminadas.

Depois todas as regiões com um ângulo com a horizontal demasiado baixo ( $\leq 40^\circ$ )

E por fim com uma proporção horizontal/vertical ( $< 0.35$ ) ou  $> 1.3$ ).

Estes parâmetros foram escolhidos por vários testes e observação dos resultados.



Figura 5. Resultado após a extração de características indesejadas

Finalmente ficamos com as regiões identificadas com maior probabilidade de serem caras onde se procedeu ao respectivo desenho da bounding box e avaliação dos resultados pelo método do Jacard.

### B. Resultados

A partir do algoritmo explicado em cima, obtivemos os seguintes resultados:



Figura 6. Resultado com uma boa identificação



Figura 7 e 8. Mau resultado

Nas figuras 6 representa uma boa identificação feita pelo algoritmo que foi sendo descrito até agora. Na identificação da pele e das bordas foram identificados mais elementos que não correspondiam à cara, no entanto estes foram eliminados posteriormente pelos métodos acima descritos.

No entanto, na figura 7 podemos ver uma má identificação da cara. Na figura 8 está representada a imagem binarizada antes de serem criadas as

bounding boxes. Nesta figura podemos ver que os métodos aplicados não foram suficientes para separar a cara do braço e de outro elemento do vestuário da pessoa em causa.

Tabela 1. Resultados da task 1

Data set	PP	TP	FP	FN	Recall (%)	Precision (%)	F-Measure (%)
Facil	41	12	29	11	52.17	29.27	37.50
Médio	78	16	62	37	30.19	20.51	24.43
Difícil	92	21	71	152	12.14	22.83	15.85

## II. DETECÇÃO E AVALIAÇÃO DE MÁSCARAS

### A. Metodologia

Neste ponto iremos descrever o algoritmo usado para fazer a deteção e avaliação do uso correcto ou incorrecto de uma máscara.

Como nos são fornecidas as coordenadas da cara das pessoas nas imagens no ground truth, então recortamos as imagens usando a função imcrop deixando assim isoladas apenas as caras.

Após o recorte foi feito um resize desta nova imagem para um tamanho fixo (400 400) que, novamente, foi escolhido por ser o tamanho da maior imagem. Isto permite que as imagens analisadas tenham o mesmo tamanho.



Figura 9. Cara recortada 400x400

Tendo as imagens individuais de cada pessoa no tamanho pretendido, convertemos novamente o espaço de cores original RGB para YCBCR da mesma maneira referida anteriormente e pelo mesmo motivo.

Podem haver valores na pele que estão fora do alcance dos valores YCBCR (ex: um reflexo de

luz na pele). Para contornar este tipo de situações foi usada a função fill seguida de um open.



Figura 10. Resultado YCBCR corrigido

Inicialmente procedeu-se a uma filtragem preliminar, na qual, por comparação de intensidades (semelhante ao realizado na secção anterior), se separou as pessoas que tinham uma máscara no rosto (independentemente de a terem bem ou mal colocada) das pessoas que não tinham máscara, por forma a facilitar a posterior análise. As imagens correspondentes a pessoas que não tinham máscara pararam de ser analisadas neste ponto.

Para as restantes imagens, procedeu-se à análise que nos permitiu aferir se a máscara estava bem colocada, mal colocada ou muito mal colocada.

Detectamos a borda superior das máscaras através da função edge.



Figura 11. Aplicação da função edge para detecção da borda da máscara

De seguida inserimos uma Bounding Box ao redor da borda superior medimos a distância entre esta e a testa da pessoa. A classificação da colocação da máscara em bem colocada, mal colocada e muito mal colocada depende dessa distância ser menor ou igual ao valor da distância de referência.

## B. Resultados

Para terminar organizamos uma matriz 1x5 que contém as avaliações de todas as imagens analisadas: “Máscara bem colocada”, “Máscara mal colocada”, “Sem máscara”, “Falsos Positivos” e “Positivos Positivos”.

CMatrix =				
63	34	5	54	5

Figura 12. Matriz de resultados da Task #2

## Referencias

- [1] Rein-Lien Hsu, Mohamed Abdel-Mottaleb and Anil K. Jain, “Face Detection in Color Images”, *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 24, nº 5, pp. 696-706, May 2002.
- [2] Amit Pal, “Multicues Face Detection in Complex Background for Frontal Faces”, *Internacional Machine Vision and Image Processing Conference*, pp. 57-62.
- [3] Peter Peer, Franc Solina, “An Automatic Human Face Detection Method”, pp. 1-9.
- [4] Q. B. Sun Hsu, W. M. Huang and J. K. Wu, “Face Detection Based on Color and Local Symmetry Information”, pp. 1-6.
- [5] Anagha S. Dhavalikar and R. K. Kulkarni, “Face Detection and Facial Expression Recognition Ststem”, pp. 6-7.
- [6] Jure Kovac, Peter Peer and Franc Solina, “Human Skin Colour Clustering for Face Detection”, pp. 1-5.
- [7] Ming-Hsuan Yang, David J. Kriegman and Narendra Ahuja, “Detecting Faces in Images: A Survey”, *IEEE Transactions on Pattern Analysis and Machine Intelligence.*, vol. 24, nº 1, pp. 34-58, January 2002.