UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS

JOÃO PAULO BROGNONI CASATI

**Método para segmentação de pele humana em imagens faciais
baseado em informações de cor e textura**

São Carlos
2013

JOÃO PAULO BROGNONI CASATI

Método para segmentação de pele humana em imagens faciais baseado em informações de cor e textura

Dissertação apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para a obtenção do título de Mestre em Ciências, Programa de Engenharia Elétrica.

Área de Concentração: Processamento de Sinais e Instrumentação.

Orientador: Prof. Dr. Evandro Luis Linhari Rodrigues.

São Carlos
2013

Trata-se da versão corrigida da dissertação. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

C336m Casati, João Paulo Brognoni
Método para segmentação de pele humana em imagens
faciais baseado em informações de cor e textura / João
Paulo Brognoni Casati; orientador Evandro Luis Linhari
Rodrigues. São Carlos, 2013.

Dissertação (Mestrado) - Programa de Pós-Graduação
em Engenharia Elétrica e Área de Concentração em
Processamento de Sinais e Instrumentação -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2013.

1. processamento digital de imagens. 2. segmentação
de pele. 3. reconhecimento de padrões. 4. textura. I.
Título.

FOLHA DE JULGAMENTO

Candidato: Bacharel **JOÃO PAULO BROGNONI CASATI**.

Título da dissertação: "Método para segmentação de pele humana em imagens faciais baseada em informações de cor e textura".

Data da defesa: 10/12/2013

Comissão Julgadora:

Resultado:

Prof. Associado **Evandro Luis Linhari Rodrigues (Orientador)**
(Escola de Engenharia de São Carlos/EESC)

Aprovado.

Prof. Dr. **Frank José Affonso**

(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP- Rio Claro)

Aprovado

Prof. Dr. **Celso Aparecido de França**

(Universidade Federal de São Carlos/UFSCar)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica e Presidente da Comissão de Pós-Graduação:

Prof. Titular **Denis Vinicius Coury**

Protocolado em 10/12/2013
Serviço de Pós-Graduação
EESC/USP

Dedico este trabalho aos meus maiores amores, minha esposa Siméia e minha filha Beatriz, que me proporcionam muitos momentos felizes e inesquecíveis, além de todo o apoio e amor que me dão nos momentos mais difíceis.

AGRADECIMENTOS

Ao Prof. Dr. Evandro Luis Linhari Rodrigues, meu orientador, por toda atenção e dedicação neste projeto de pesquisa científica.

Aos membros da banca de qualificação, Profs. Drs.: Frank José Affonso e Celso Aparecido de França; pela participação e dedicação nas sugestões de correções deste trabalho.

Ao corpo docente que ministrou as minhas disciplinas, Profs. Drs.: Adilson Gonzaga, Ivan Nunes da Silva, Marcelo Andrade da Costa Vieira e Maria Stela Veludo de Paiva; pelos conhecimentos e conselhos acadêmicos, que somados resultaram no profissional que sou.

Aos funcionários da USP de São Carlos, sempre atenciosos e dispostos a ajudar em qualquer tipo de problema, em especial à Marisa Helena R. Vicente Fortulan, Leonardo Erivelto Soares de Oliveira, Roseli Ribeiro e Rosane A. Aranda.

Aos Profs. Drs.: Adilson Gonzaga e Osvaldo Severino Júnior, pela disponibilização das imagens de *ground truths* utilizadas neste trabalho.

À minha filha Beatriz Nunes Casati e à minha esposa Siméia Rafaela Nunes Casati por todo apoio, pelo incentivo e por serem as razões da minha dedicação.

Aos meus pais, Paulo Carvalho Casati e Cassandra Marta Alves Brognoni Casati pelo amor, carinho e instrução que foram fundamentais na formação do meu caráter.

Por último, mas não menos importantes, a todos os meus amigos e colegas pessoais, profissionais e da pós-graduação, que sempre estiveram disponíveis a me ajudar em qualquer tipo de problema, os quais sabem que não é necessário listar os seus nomes aqui, porque verdadeiramente conhecem a sua importância para mim.

RESUMO

CASATI, J. P. B. **Método para segmentação de pele humana em imagens faciais baseado em informações de cor e textura**. 2013. 107 f. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2013.

A segmentação de pele em imagens é um importante processo para uma vasta gama de aplicações, como detecção e rastreamento de faces, reconhecimento de gestos, computação forense, entre outros. Um dos maiores problemas encontrados neste tipo de aplicação é a presença de objetos que possuem cor de pele nas imagens, mas não fazem parte de segmentos reais de pele, sendo muitas vezes erroneamente classificados como pele. A fim de reduzir a frequência destes falsos positivos, é apresentado neste trabalho um método de segmentação de pele humana em imagens faciais que possui duas diferentes etapas que reduzem a quantidade de falsos positivos do processo sem que se percam quantidades significantes de verdadeiros positivos. Estas duas etapas são chamadas de FPAR (*False Positive Area Reduction*) e aplicação de textura. A primeira visa remover segmentos não contínuos classificados como pele e a segunda aborda a aplicação de textura nas imagens, removendo áreas em que a textura não se assemelha à textura de pele humana. Para isto, foi desenvolvido o banco de imagens SFA (*Skin of FERET and AR*), constituído de imagens originais dos bancos de faces FERET e AR, seus respectivos *ground truths* de segmentação de pele e amostras de pele e não pele extraídas das imagens originais. O método apresentado neste trabalho apresenta resultados promissores atingindo até 46,9% de redução de falsos positivos sem que a acurácia aferida tenha redução significativa (apenas 1,8%). Este trabalho tem como contribuições o método desenvolvido e o banco de imagens SFA que fica disponível online para download pela comunidade científica.

Palavras-chave: processamento digital de imagens, segmentação de pele, reconhecimento de padrões, textura.

ABSTRACT

CASATI, J. P. B. **A method for human skin segmentation in facial images based on color and texture features**. 2013. 107 f. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2013.

Skin segmentation is an important process for many kinds of application like face detection, face tracking, gesture recognition, forensic computing and others. One of the main problems found in this kind of application is the presence of objects which have skin color but are not part of actual skin segments, being wrongly classified as skin. Aiming to reduce the frequency of these false positives, this work presents a method of human skin segmentation in facial images which has two different steps that reduces the false positives without losing significant areas of true positives. These two steps are called FPAR (False Positive Area Reduction) and texture application. The first one removes segments classified as skin which are not continuous and the second one is an analysis of the image texture, removing areas which the texture is not alike human skin texture. To achieve this, the SFA (Skin of FERET and AR) image database was developed, constituted of original images retrieved from AR and FERET face databases, their respective ground truths of skin segmentation and skin and non-skin samples retrieved from the original images. The method presented in this work shows promising results, reaching up to 46.9% of false positive reduction without significant reduction of the accuracy (1.8%). This work has as contributions the developed method and the SFA database, which is available for download for scientific community.

Keywords: digital image processing, skin segmentation, pattern recognition, texture.

LISTA DE FIGURAS

Figura 1 - Representação do cubo RGB.	28
Figura 2 - Exemplo de formação de imagem colorida em RGB, onde (a) é a imagem de intensidade do canal R, (b) é a imagem de intensidade do canal G, (c) é a imagem de intensidade do canal B e (d) é a imagem resultante da junção dos canais R, G e B.	28
Figura 3 - Representação do espaço de cores HSV.	30
Figura 4 - Etapas de pintura artística em tela.	31
Figura 5 - Imagens faciais quantizadas pelo misturograma.	32
Figura 6 - Exemplo de aplicação do LFP.	34
Figura 7 - Neurônio artificial.	35
Figura 8 - Arquitetura da MLP.	36
Figura 9 - Topologia de rede MLP (a) e espaço amostral classificado (b).	37
Figura 10 - Esboço da face gerado pelo misturograma.	38
Figura 11 - Pele do indivíduo segmentada da imagem.	40
Figura 12 - (a) imagem original e (b) imagem com compensação de iluminação.	40
Figura 13 - Resultado final, onde (a) é a região de face binarizada e (b) é a região de pele segmentada.	41
Figura 14 - Processo de detecção de pele.	42
Figura 15 - Resultados de Jusoh et al. (2010).	42
Figura 16 - Classes de imagens do banco AR.	46
Figura 17 - Imagem original pré-processada (a) e imagem quantizada (b).	48
Figura 18 - Imagem pré-processada (a) e esboço da imagem gerado pelo misturograma (b).	48
Figura 19 - Imagens de face quantizadas pelo misturograma e seus respectivos histogramas de cores abaixo.	49
Figura 20 - Histogramas das imagens acrescido dos histogramas do esboço.	50
Figura 21 - As figuras (a) e (b) são exemplos de imagens originais do banco AR e as figuras (c) e (d) são exemplos de imagens originais do banco FERET.	54
Figura 22 - Gráfico de distribuição da quantidade de imagens utilizadas de cada banco.	54
Figura 23 - Imagens originais seguidas de seus respectivos Ground Truths.	55
Figura 24 - Exemplo de criação das amostras.	57
Figura 25 - Estrutura de pastas do SFA.	57
Figura 26 - Topologias da RNA para validação das amostras.	58
Figura 27 - (a) Imagem original do banco SFA e (b) imagem do GT binarizada.	60
Figura 28 - Topologias de RNA utilizadas para segmentação de pele nas imagens.	62
Figura 29 - Diagrama de processos do FPAR.	63
Figura 30 - Diagrama dos processos para análise de textura.	65
Figura 31 - Imagem original do banco SFA (a) e imagem resultante da aplicação do LFP (b).	66
Figura 32 - (a) Imagem original em escala de cinza, (b) IMG_DIST com amostras de 5x5, (c) IMG_DIST com amostras de 11x11, (d) IMG_DIST com amostras de 15x15, (e) IMG_DIST com amostras de 25x25, (f) IMG_DIST com amostras de 35x35.	68
Figura 33 - Comportamento da acurácia em relação ao <i>threshold</i> aplicado para ambos os bancos (UCI e SFA).	74

Figura 34 - Comparação entre a segmentação utilizando SFA e UCI. A primeira couna (a) contém as imagens originais, a segunda coluna (b) contém os GT, a terceira coluna (c) contém as imagens segmentadas utilizando amostras do SFA e a quarta coluna (d) contém as imagens segmentadas utilizando as amostras do UCI.	75
Figura 35 - Gráfico de acurácia dos espaços de cores.....	78
Figura 36 - Gráfico do comportamento da acurácia com a aplicação do FPAR.	80
Figura 37 - Gráfico do comportamento dos FP com a aplicação do FPAR e parâmetro de 15%.	82
Figura 38 - Comparativo de FP entre RNA e FPAR.	82
Figura 39 - Acurácia em relação aos valores de β para textura aplicada em IRNA..	84
Figura 40 – Falsos positivos em relação aos valores de β para textura aplicada em IRNA.	84
Figura 41 - Acurácia em relação aos valores de β para textura aplicada em IFPAR.	86
Figura 42 - Falsos positivos em relação aos valores de β para textura aplicada em IFPAR.	86
Figura 43 - Resultado do método proposto para a imagem: img (20).jpg.	88
Figura 44 - Resultado do método proposto para a imagem: img (22).jpg.	89
Figura 45 - Comparativo das acurácias das aplicações.	89
Figura 46 - Comparativo dos falsos positivos das aplicações.....	90

LISTA DE TABELAS

Tabela 1 - Resultados de Wang, Chang e Liu (2009).	41
Tabela 2 - Resultados de Wang, Zhang e Yao (2011).	43
Tabela 3 - Vetor de características das imagens quantizadas.	50
Tabela 4 - Vetor de características de imagens quantizadas acrescido de informações de cores do esboço da face.	51
Tabela 5 - Vetor de características de 7 posições.	52
Tabela 6 - Quantidade de Imagens do SFA.	56
Tabela 7 - Acertos da abordagem com vetor de 8 posições.	70
Tabela 8 - Acertos da abordagem com vetor de 10 posições.	70
Tabela 9 - Acertos da abordagem com vetor de 7 posições.	71
Tabela 10 - Medidas de velocidade de processamento em segundos.	71
Tabela 11 - Resultados da validação para grupo de amostras de teste da RNA.	72
Tabela 12 - Resultados de acurácia para a segmentação utilizando SFA e UCI.	73
Tabela 13 - Valores de correlação por espaços de cores e topologia da RNA.	76
Tabela 14 - Acurácia para segmentação das imagens por espaço de cores.	77
Tabela 15 - Resultados de FP para a aplicação de RNA no espaço de cores RGB.	78
Tabela 16 - Valores de acurácia com aplicação do FPAR.	79
Tabela 17 - Resultados de FP para a aplicação do FPAR.	81
Tabela 18 - Resultados da aplicação de textura nas imagens IRNA.	83
Tabela 19 - Resultados ótimos da aplicação de textura nas imagens IRNA.	85
Tabela 20 - Resultados da aplicação de textura nas imagens IFPAR.	85
Tabela 21 - Resultados ótimos da aplicação de textura nas imagens IFPAR.	87

LISTA DE ABREVIATURAS E SIGLAS

CBIR	<i>Content Based Image Retrieval</i>
FP	<i>Falsos Positivos</i>
FPAR	<i>False Positive Area Reduction</i>
IFPAR	<i>Imagem resultante do FPAR</i>
IRNA	<i>Imagem resultante da RNA</i>
LFP	<i>Local Fuzzy Pattern</i>
MATLAB	<i>Matrix Laboratory</i>
MLP	<i>Multi-Layer Perceptron</i>
r	Coeficiente de correlação [-1 1]
$r_{\max}$	Maior valor de correlação
RNA	Redes Neurais Artificiais
SFA	<i>Skin of FERET and AR</i>
VP	<i>Verdadeiros Positivos</i>

SUMÁRIO

1.	Introdução	23
1.1.	Objetivo	24
1.2.	Justificativa	25
1.3.	Organização do Trabalho	25
1.4.	Publicações Relacionadas.....	26
2.	Fundamentação Teórica e Trabalhos Relacionados	27
2.1.	Imagens Coloridas e Espaços de Cores.....	27
2.1.1.	Misturograma	31
2.2.	Textura e o LFP	33
2.3.	Redes Neurais Artificiais.....	34
2.3.1.	Redes Perceptron Multicamadas (MLP)	36
2.4.	Trabalhos Relacionados	37
2.4.1.	Cores e Reconhecimento Facial	38
2.4.2.	Segmentação de Pele.....	39
3.	Método Proposto	45
3.1.	Reconhecimento Facial	45
3.1.1.	Pré-Processamento	47
3.1.2.	Quantização de Cores e Esboço	47
3.1.3.	Extração de Características	49
3.1.4.	Classificação.....	52
3.2.	Banco de Imagens SFA.....	53
3.2.1.	Construção do SFA.....	53
3.2.2.	Validação com Classificação das Amostras.....	58
3.2.3.	Validação com Segmentação de Imagens.....	59
3.3.	Segmentação da Pele	61
3.3.1.	Segmentação por RNA em Diferentes Espaços de Cores.....	61
3.3.2.	Aplicação do FPAR.....	63
3.3.3.	Aplicação de Textura	64
4.	Resultados	69
4.1.	Reconhecimento Facial	69
4.2.	Banco de Imagens SFA.....	72
4.2.1.	Validação com Classificação das Amostras.....	72
4.2.2.	Validação com Segmentação de Imagens.....	73
4.3.	Segmentação da Pele	75
4.3.1.	Segmentação por RNA em Diferentes Espaços de Cores.....	76
4.3.2.	Aplicação do FPAR.....	78
4.3.3.	Aplicação de Textura	83
5.	Conclusões	91
5.1.	Contribuições.....	92
5.2.	Sugestões para Trabalhos Futuros	92
	Referências Bibliográficas	93
	Apêndice A – Exemplos de Imagens do SFA.....	99
	Apêndice B – Código-fonte: transformação do RGB para RGB normalizado..	103
	Apêndice C – Código-fonte: Procedimentos do Método Proposto.....	105

1. INTRODUÇÃO

Em visão computacional a detecção de pele humana em imagens digitais é utilizada em diversos tipos de aplicação, sendo algumas delas: reconhecimento facial (CHAVES-GONZÁLEZ et al., 2010), detecção de faces (JIN et al, 2007), rastreamento de faces (SARKAR; BAKSHI; SA, 2012), computação forense, sistemas de avaliação de cirurgias plásticas (BAKHSHALI; SHAMSI; GOLZARFAR, 2012), entre outras.

Falcão (2013)¹ afirma que na área de computação forense, há uma vasta gama de aplicações de detecção de pele, como por exemplo, a detecção de nudez em imagens para se rastrear imagens de pedofilia, a detecção de faces em imagens para se filtrar imagens que possam conter contatos e relações suspeitas, entre muitas outras (informação verbal).

A detecção de pele em imagens digitais pode ser dividida em abordagens baseada em pixels (VEZHNEVETS; SAZONOV; ANDREEVA, 2003) e abordagem baseada em regiões (JIANG; YAO; JIANG, 2007). A abordagem baseada em pixels considera os valores de cor de cada pixel e classifica-o como pele ou não pele. Na abordagem baseada em região, outras técnicas podem ser aplicadas, a principal delas é a análise de textura (NG; PUN, 2011). Um dos problemas mais recorrentes em sistemas de segmentação de pele é a alta frequência de falsos positivos. Isto ocorre devido a objetos pertencentes a imagem que tenham cor semelhante à cor de pele humana. Estes objetos costumam ser classificados erroneamente como segmentos de pele (SIQUEIRA; SCHWARTZ; PEDRINI, 2013).

A grande maioria das aplicações em detecção de pele humana em imagens é feita utilizando modelagem de cores, mas com o poder computacional atual muitas aplicações utilizam de informações de textura e informações geométricas para que a detecção seja mais robusta e alcance melhores resultados (KAKUMANU; MAKROGIANNIS; BOURBAKIS, 2007). As aplicações de detecção baseadas em

¹ Informação fornecida por Alexandre Xavier Falcão em palestra no IX Workshop de Visão Computacional no Rio de Janeiro-RJ, em 2013.

cores normalmente variam de espaço de cores utilizado. Estes espaços de cores são profundamente estudados e os resultados alcançados por cada espaço de cor pode variar consideravelmente (CHAVES-GONZÁLEZ et al., 2010).

Devido ao problema da ocorrência de falsos positivos, este trabalho apresenta um método de segmentação de pele humana em imagens digitais que visa reduzir a quantidade destes falsos positivos erroneamente classificados como pele, sem que a acurácia obtida na segmentação tenha redução significativa. Primeiramente, a pele humana é detectada na imagem com um processamento baseado em pixels, feito por uma Rede Neural Artificial (RNA). Posteriormente, as duas etapas de redução de falsos positivos são aplicadas nas imagens já segmentadas, visando à redução dos falsos positivos. Para confirmar que as cores são bons descritores de faces humanas, a primeira etapa do método do trabalho é uma aplicação de reconhecimento facial utilizando apenas cores como descritores. A segunda etapa é a criação de um banco de imagens completo para a aplicação de segmentação de pele, com *ground truths*, amostras de pele e não pele. Estas amostras são utilizadas no treino da RNA para que as imagens originais sejam segmentadas.

Quanto às etapas de redução de falsos positivos, a primeira aplicação é o FPAR (*False Positive Area Reduction*), que consiste em remover pequenas áreas não contínuas da imagem que tenham sido classificadas como pele. A segunda etapa é a aplicação de textura nas imagens para que as áreas em que a textura não se assemelhe à textura de pele possam ser removidas. Neste procedimento, é utilizada a análise de micropadrões baseada em números *fuzzy* (LFP), será apresentada no item 2.2.

1.1. Objetivo

O principal objetivo deste trabalho foi desenvolver um método de detecção de pele humana em imagens de face que seja direcionado para a minimização de falsos positivos. Para isto, duas etapas de redução de falsos positivos foram desenvolvidas e aplicadas nas imagens segmentadas pela RNA. Ao aplicar as

etapas de redução de falsos positivos, procurou-se manter a acurácia da segmentação em níveis altos (mais que 90%).

1.2. Justificativa

Métodos de detecção de pele são muito utilizados em diversos tipos de problema. Porém, as cores de pele podem ser confundidas com as cores de outros objetos em uma imagem. Assim, este trabalho aplica informações de textura visando minimizar a ocorrência de falsos positivos, ou seja, remover pixels que foram detectados como pele por sua semelhança com a cor da pele humana, mas que possuem a textura diferente. Com isto, podem-se aperfeiçoar os algoritmos e aplicações que possuem esta característica como variável de entrada.

1.3. Organização do Trabalho

Este trabalho está organizado de forma a introduzir o leitor ao campo de pesquisa no Capítulo 1, mostrando quais as necessidades de se pesquisar métodos de detecção de pele e também exibir as publicações do autor durante o período de desenvolvimento do mesmo. No Capítulo 2, são exibidas de forma breve algumas das tecnologias utilizadas no desenvolvimento deste trabalho, assim como alguns trabalhos anteriores, relacionados direta ou indiretamente a este, e que são de fundamental importância para seu desenvolvimento.

No Capítulo 3 é apresentado de forma completa o método desenvolvido neste trabalho em todas as suas etapas. Posteriormente, no Capítulo 4, são apresentados os resultados obtidos com este trabalho, assim como a discussão dos mesmos.

Por fim, no Capítulo 5 é apresentada a conclusão do autor baseada nos resultados obtidos assim como suas contribuições para a comunidade científica e algumas opções já identificadas que possibilitam a continuação da pesquisa.

1.4. Publicações Relacionadas

Neste item são relacionadas três publicações do autor relacionadas ao trabalho desenvolvido.

A primeira publicação, intitulada “Comparação de Técnicas de Reconhecimento Facial Utilizando Histogramas de Cores Quantizadas pelo Misturograma”, foi submetida e aceita no VIII Workshop de Visão Computacional de 2012, realizado na cidade de Goiânia-GO (CASATI; RODRIGUES, 2012). Trata-se de uma aplicação de reconhecimento facial utilizando apenas cores quantizadas. Esta aplicação visa demonstrar a importância que as cores exercem no contexto da face humana. O detalhamento do desenvolvimento desta aplicação encontra-se no item 3.1 e os resultados obtidos encontram-se no item 4.1.

Na segunda publicação, intitulada “*Analysis of Polynomial Behavior of the C3 Cervical Concavity to Bone Age Estimation using Artificial Neural Networks*” (MORAES; CASATI; RODRIGUES, 2013), que foi submetida e aceita no “*Biosignals and Biorobotics Conference*” (BRC) de 2013 realizado na cidade do Rio de Janeiro-RJ, foi apresentado um método de estimação de idade óssea utilizando uma característica inédita na literatura, o polinômio extraído da concavidade da vértebra C3. No caso deste trabalho, foi desenvolvido também um método de avaliação de topologia de redes neurais artificiais que está diretamente ligado a este trabalho, sendo utilizado nos itens 3.2.2, 3.2.3 e 3.3.1.

A terceira publicação, intitulada “*SFA: A Human Skin Image Database based on FERET and AR Facial Images*”, foi submetida e aceita no IX Workshop de Visão Computacional de 2013, realizado na cidade do Rio de Janeiro-RJ (CASATI; MORAES; RODRIGUES, 2013). Este artigo apresenta a criação e validação do banco de imagens SFA. Este banco de imagens é a base para o desenvolvimento do método proposto neste trabalho. Os detalhes do desenvolvimento e validação deste banco de imagens são apresentados no item 3.2.

Estas publicações corroboram a importância e algumas contribuições científicas do trabalho desenvolvido. O próximo capítulo apresenta, de forma resumida, os tópicos que serão abordados no decorrer deste trabalho.

2. FUNDAMENTAÇÃO TEÓRICA E TRABALHOS RELACIONADOS

Neste capítulo é dada uma introdução aos tópicos que são utilizados e discutidos ao longo do trabalho.

2.1. Imagens Coloridas e Espaços de Cores

As cores são poderosos descritores de imagens digitais, podendo muitas vezes simplificar o reconhecimento de objetos, a segmentação da imagem ou até mesmo sua interpretação (GONZALEZ; WOODS, 2008). O olho humano identifica as cores por meio de seus cones, sensíveis às cores: vermelho, verde e azul. As demais cores são formadas por combinações destas três cores primárias (GONZALEZ; WOODS, 2008; KOSCHAN; ABIDI, 2008).

As imagens digitais coloridas têm suas cores representadas por espaços de cores (chamados também de modelos de cores ou sistemas de cores). São especificações de sistemas de coordenadas, onde cada cor é representada por um único ponto, ou seja, representações matemáticas geométricas das cores.

Os espaços de cores utilizados neste trabalho são o RGB, o YCbCr, o HSV e o RGB Normalizado (NormRGB). O espaço de cores RGB é baseado em como o olho humano interpreta as cores. Seus três canais de cores (três dimensões do cubo RGB) são: vermelho (*Red*), verde (*Green*) e azul (*Blue*). O cubo RGB está representado na Figura 1.

A representação das cores é feita em 3 dimensões, uma imagem possui 3 subimagens (canais de cor) em tons de cinza (grau de luminosidade). Cada canal de duas dimensões representa uma cor. Na Figura 2 é apresentado um exemplo dos três canais de cor se fundindo e gerando uma imagem colorida no espaço de cores RGB.

Figura 1 - Representação do cubo RGB.

Figura 2 - Exemplo de formação de imagem colorida em RGB, onde (a) é a imagem de intensidade do canal R, (b) é a imagem de intensidade do canal G, (c) é a imagem de intensidade do canal B e (d) é a imagem resultante da junção dos canais R, G e B.

Outro espaço de cores utilizado neste trabalho é o YCbCr. Este espaço de cores foi criado para que as emissões dos sistemas de televisão em cores fossem compatíveis também com os receptores preto e branco. Os sistemas SECAM (*Séquence Electronique Couleur Mémoire*) e PAL (*Phase Alternating Line*) utilizam o espaço de cores YCbCr (SEVERINO JUNIOR, 2009).

O YCbCr separa o sinal de cores RGB em um sinal Y (grau de luminosidade) e dois sinais de cores: Cb (azul) e Cr (vermelho). (SPROSON, 1983).

A transformação do espaço de cores RGB para o YCbCr é dada pelas equações 1, 2 e 3 (SEVERINO JUNIOR, 2009), sendo os valores R, G e B advindos do espaço de cores RGB.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = M \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = M^{-1} \begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} \quad (2)$$

$$M = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ -0,169 & -0,331 & 0,5 \\ 0,5 & -0,419 & -0,0813 \end{bmatrix} \quad (3)$$

O espaço de cores HSV é amplamente utilizado para segmentação de pele em imagens (SIQUEIRA; SCHWARTZ; PEDRINI, 2013; KAKUMANU; MAKROGIANNIS; BOURBAKIS, 2007).

Os três eixos deste espaço de cores são: *Hue* (H), que significa matiz, *Saturation* (S) que significa saturação e *Value* (V) que significa valor. O HSV é representado por uma pirâmide, representada na Figura 3.

O eixo H representa a cor, ou seja, a tonalidade. O eixo S representa a saturação, que é o quanto esta cor é pura, ou seja, quanto mais próximo de zero, mais parecida com tom de cinza a cor é, e quanto maior seu valor, mais pura é a aparência desta cor. O eixo V representa o valor, ou seja, a intensidade da cor (brilho) (KOSCHAN; ABIDI, 2008).

Figura 3 - Representação do espaço de cores HSV.

A transformação do espaço de cores RGB para o HSV dá-se pelas equações 4 – 6.

$$H = \begin{cases} 60 * \frac{G-B}{MAX-MIN} + 0 \rightarrow se (MAX = R) e (G \geq B) \\ 60 * \frac{G-B}{MAX-MIN} + 360 \rightarrow se (MAX = R) e (G < B) \\ 60 * \frac{B-R}{MAX-MIN} + 120 \rightarrow se (MAX = G) \\ 60 * \frac{R-G}{MAX-MIN} + 240 \rightarrow se (MAX = B) \end{cases} \quad (4)$$

$$S = \frac{MAX-MIN}{MAX} \quad (5)$$

$$V = MAX \quad (6)$$

Sendo MAX e MIN os valores de máximo e mínimo dos valores R, G e B, advindos do espaço de cores RGB.

O espaço de cores RGB normalizado (NormRGB) é uma variação do espaço de cores RGB que visa reduzir a dependência de iluminação das imagens. A normalização impõe que a soma dos valores R, G e B sejam iguais a um, como demonstrado na equação 7 (KAKUMANU; MAKROGIANNIS; BOURBAKIS, 2007).

$$R + G + B = 1 \quad (7)$$

Com esta normalização, consegue-se também a redução na dimensão, pois apenas dois canais são relevantes, o terceiro pode ser obtido a partir deles. O código-fonte da transformação do espaço de cores RGB para o NormRGB é exibido no Apêndice B – Código-fonte: transformação do RGB para RGB normalizado. Este código é uma implementação do algoritmo de transformação para MATLAB.

A seguir, é apresentada a técnica do misturograma, utilizado na aplicação apresentada no item 3.1.

2.1.1. Misturograma

O misturograma é um método de quantização de cores baseado na forma em que os artistas compõem suas telas, sendo cada cor formada por diversas camadas de diferentes cores. O espaço de cores RGB é utilizado para esta quantização. As imagens têm suas cores quantizadas a partir de processos matemáticos e resultam em uma imagem de oito cores (SEVERINO JUNIOR; GONZAGA, 2005). Além da significativa redução no número de cores das imagens, suas diferentes camadas permitem uma segmentação da imagem, como a geração do esboço de faces e segmentação de pele humana (SEVERINO JUNIOR; GONZAGA, 2006).

Figura 4 - Etapas de pintura artística em tela.

Fonte: RIBEIRO², 2005 apud SEVERINO JUNIOR, 2009, p.60.

² RIBEIRO, S. E. M. **Ilustrações lápis, nanquim e cores**. Silvio Ribeiro, 2005.

A quantização pelo misturograma gera oito camadas de oito possíveis cores cada, e a imagem final é resultado da sobreposição destas camadas multiplicadas por seus devidos pesos. As cores geradas pelo misturograma são: preto, azul, verde, ciano, vermelho, magenta, amarelo e branco. Este método de quantização de cores proporciona uma gama de possibilidades, como pode-se observar nos outros trabalhos analisados no item 2.4.

Exemplos de imagens de face quantizadas pelo misturograma são exibidos na Figura 5. Observa-se grande redução na quantidade de cores das imagens, porém, as principais informações espaciais da imagem não são perdidas. Isto ocorre devido ao seu poder de segmentação da imagem.

Figura 5 - Imagens faciais quantizadas pelo misturograma.

Sustentado por diversas publicações, como se pode observar no item 2.4, o misturograma se mostra uma poderosa ferramenta de segmentação e quantização de cores.

No próximo item é apresentado um breve embasamento teórico sobre texturas e o LFP, método de análise de micropadrões utilizado no procedimento de aplicação de textura deste trabalho.

2.2. Textura e o LFP

O termo textura não possui uma definição em termos gerais, normalmente esta definição é específica para o tipo de aplicação em questão. Portanto, diversas definições são encontradas na literatura (HARALICK; SHANMUGAN; DINSTEIN, 1973; BIMBO, 1999).

O LFP (*Local Fuzzy Pattern*) é um método de análise de micropadrões de textura em imagens digitais que se baseia em números *fuzzy*. Estes micropadrões são estruturas de pixels em níveis de cinza que descrevem um contexto espacial da imagem (VIEIRA, 2013).

Sustentado por diversas publicações (VIEIRA et al., 2012a; VIEIRA et al., 2012b; VIEIRA et al., 2012c), o método obteve resultados promissores e por se tratar de uma técnica atual encontra-se em constante desenvolvimento.

O método LFP é apresentado em (VIEIRA, 2013) e possui aplicação utilizando a função sigmoide (LFP-s), a função triangular (LFP-t) e a função gaussiana (LFP-g). Para a aplicação do LFP neste trabalho, utilizou-se a função sigmoide, que obtém melhores resultados para análise de micropadrões em imagens.

O LFP-s é o procedimento utilizado neste trabalho para a aplicação de redução de falsos positivos com textura nas imagens de segmentação de pele. É escolhida esta variação do LFP por ter apresentado os melhores resultados na análise de texturas (VIEIRA, 2013).

A Figura 6 mostra um exemplo da aplicação do LFP na descrição de um micro padrão de textura.

Figura 6 - Exemplo de aplicação do LFP.

Fonte: Vieira (2013).

O pixel da imagem (nível de cinza) é a entrada para a função sigmoide, sendo o valor do pixel central como centro. A saída desta função é o grau de pertinência do pixel (μ), gerando assim a matriz de pertinência. A matriz de pertinência é multiplicada ponto a ponto por uma matriz de pesos previamente definida. A média ponderada arredondada dos pixels desta matriz é atribuída ao pixel central. Este procedimento é executado percorrendo toda a imagem, gerando uma imagem de saída LFP. Para a análise dos padrões, usa-se o histograma local da área a ser analisada. Esta análise é mais bem detalhada no item 3.3.3.

2.3. Redes Neurais Artificiais

Redes neurais artificiais são modelos computacionais baseados no comportamento e estrutura do cérebro humano. Composta por um ou mais neurônios artificiais, dispostos em uma ou mais camadas, que trocam ou não sinais entre si, as redes são dotadas de inteligência artificial, ou seja, aprendem com um conjunto de sinais de entrada (amostras de treinamento) e procuram generalizar este aprendizado para todo o universo de sinais de entrada (SILVA; SPATTI; FLAUZINO, 2010).

A composição de um neurônio é exibida na Figura 4, onde:

- $X_{1..n}$ – sinais de entrada;
- $W_{1..n}$ – pesos sinápticos;

- θ – limiar de ativação;
- Σ – combinador linear;
- u – potencial de ativação;
- g – função de ativação;
- y – sinal de saída.

Figura 7 - Neurônio artificial.

Fonte: Silva, Spatti e Flauzino (2010).

As redes neurais artificiais datam de mais de 50 anos desde a primeira publicação a seu respeito, a arquitetura do *Perceptron* Simples. Porém, só a partir dos anos 1990 foram amplamente estudadas, e vieram a se tornar uma poderosa ferramenta para o desenvolvimento de sistemas inteligentes e resolução de problemas utilizando inteligência artificial (SILVA; SPATTI; FLAUZINO, 2010).

Muitas arquiteturas de rede foram criadas entre os anos de 1970 e 1980, como: LVQ (*Learning Vector Quantization*), Hopfield, Kohonen, ART (*Adaptive Resonance Theory*), entre outras. Estas arquiteturas incentivaram a retomada dos estudos e pesquisas sobre redes neurais artificiais, que estavam estagnadas há muitos anos. Outra grande contribuição para que as redes neurais voltassem à tona foi a descoberta do algoritmo *backpropagation*, utilizado para treinamento de redes neurais de mais de uma camada, como a RBF (*Radial Basis Function*) e a MLP (*Multi-Layer Perceptron*), esta última sendo a mais utilizada de todas, podendo resolver problemas de aproximação universal de funções, reconhecimento de padrões, identificação e controle de processos, otimização de sistemas e previsão de séries temporais (HAYKIN, 1999).

A decisão de qual arquitetura e topologia de redes neurais utilizar é muito importante, pois cada problema necessita de um tipo de solução, e cada arquitetura de rede proporciona uma ferramenta diferente para tal.

No item seguinte é dada uma breve introdução sobre a principal arquitetura de rede neural aplicável ao reconhecimento de padrões. Esta arquitetura de rede também podem resolver outros tipos de problema, como previsão em sistemas variantes no tempo e aproximação de funções, porém esta introdução limita-se a dar uma breve descrição sobre o seu uso em classificação de padrões.

2.3.1. Redes Perceptron Multicamadas (MLP)

A rede MLP (do inglês *Multi-Layer Perceptron*) multicamadas é derivada da rede *perceptron* simples, que contém apenas uma camada de um neurônio, portanto, só consegue classificar entre 2 diferentes padrões, que devem estar linearmente separados no espaço amostral (SILVA; SPATTI; FLAUZINO, 2010).

Uma das características da rede MLP é conter ao menos uma camada escondida de neurônios (camada intermediária). Um exemplo da topologia da MLP é exibido na Figura 8.

Figura 8 - Arquitetura da MLP.

Fonte: Silva, Spatti e Flauzino (2010).

O que permitiu o treinamento de redes neurais multicamadas foi a descoberta do algoritmo *backpropagation*. Basicamente, este algoritmo executa a entrada da rede, mede o erro e faz o caminho de volta alterando os pesos sinápticos. Este processo é executado até que o erro esteja dentro da margem previamente definida pelo usuário.

Diferentemente do *perceptron* simples, as amostras a serem classificadas não precisam necessariamente ser linearmente separáveis. Um exemplo da versatilidade na classificação de padrões é exemplificado na Figura 9, onde um espaço amostral é dividido conforme a topologia da rede neural.

Figura 9 - Topologia de rede MLP (a) e espaço amostral classificado (b).

Fonte: adaptado de Silva, Spatti e Flauzino (2010).

A escolha da topologia de rede neural a ser utilizada é essencial para uma eficiente classificação dos padrões. O espaço amostral deve ser classificado de forma a generalizar o problema, o que torna a rede neural eficiente na aplicação da inteligência artificial (PERETTO, 1992).

2.4. Trabalhos Relacionados

Embora existam muitos trabalhos que possam ter relação com este na literatura, alguns dos mais semelhantes e relevantes são brevemente descritos nesta seção, organizados por tema e data de publicação.

2.4.1. Cores e Reconhecimento Facial

Uma abordagem de reconhecimento facial baseada em cores quantizadas pelo misturograma foi apresentada por Severino Junior e Gonzaga (2005). Neste método, a imagem é reduzida a 8 cores e seus histogramas são utilizados para a comparação entre as imagens.

Métodos de computação evolutiva foram utilizados para a extração de características baseadas em cores para o reconhecimento facial. As imagens são transformadas de RGB para dois outros espaços de cores e testadas. Observou-se que os resultados utilizando o modelo de computação evolutiva melhoram o desempenho dos algoritmos (SHIH; LIU, 2005).

Segundo Severino Junior e Gonzaga (2006), o esboço da imagem facial e a segmentação da pele humana podem ser obtidos por meio do misturograma, através de suas diferentes camadas. O esboço da face pode ser obtido considerando apenas os pixels de cor preta da camada 7. A Figura 10 exibe o esboço da face gerado pelo misturograma.

Figura 10 - Esboço da face gerado pelo misturograma.

Alguns dos principais métodos de extração de características baseados em cores são comparados por Penatti e Torres (2008) para imagens aleatórias, com o objetivo de se apresentar uma imagem como entrada e buscar o conjunto de imagens mais relevantes (similares), simulando um ambiente de busca de imagens pela internet. A avaliação dos métodos é baseada em sua velocidade de extração das características, tamanho do vetor de características gerado e velocidade de comparação dos vetores (classificação).

No artigo de Silva e Gonzaga (2009), um sistema CBIR é construído a fim de se fazer o reconhecimento facial por meio dos esboços de faces gerados pelo misturograma e retratos falados. Estes retratos são construídos a partir de softwares como Corel Draw, Photoshop e FlashFace e Photocomposer Plus. A referência para a criação desses retratos é a imagem original do indivíduo. Os retratos são armazenados em uma base de imagens e posteriormente são comparados com o esboço de imagens de indivíduos.

Jafri e Arabnia (2009) publicaram uma pesquisa (*survey*) de reconhecimento facial muito completo, abordando as dificuldades e diversas técnicas utilizadas em diferentes tipos de problemas em reconhecimento facial.

No artigo escrito por Celebi, Wen e Chen (2011) é feito um comparativo entre o uso dos algoritmos de clusterização *Fuzzy C-Means* e *Hard C-Means* aplicados na quantização de cores de imagens. Concluiu-se com o artigo publicado que o método *Fuzzy C-Means* é mais lento e não oferece resultados melhores que o *Hard C-Means*.

Choi, Ro e Plataniotis (2011) utilizam de informações de cores para desenvolver um método de reconhecimento facial. As imagens são transformadas em vários espaços de cores diferentes e suas características são extraídas separadamente. Após isto, uma fusão das características extraídas ocorre, as características recebem pesos diferentes dependendo do quanto contribuem para um bom reconhecimento facial, e por fim, as imagens são classificadas. Vários bancos de imagens foram utilizados no desenvolvimento deste trabalho, alcançando resultados expressivos na melhora da seleção das características de cor das imagens.

QiaoQiao e Yuanyuan (2012) utilizam um método de quantização de cores auto-organizável e utilizam o histograma da imagem quantizada como característica para um sistema CBIR, muito parecido com a abordagem utilizada neste trabalho, porém, não foi testado em imagens de face.

2.4.2. Segmentação de Pele

Severino Jr. e Gonzaga (2006) aplicam o misturograma na segmentação de pele. A imagem é transformada do espaço de cores RGB para HSV. Trabalhando no

espaço de cores HSV, os canais S e V são normalizados. Já com os canais normalizados, a imagem é novamente transformada para o espaço de cores RGB. Por fim, a imagem é quantizada pelo misturograma a fim de segmentar a região de pele. A Figura 11 mostra a imagem original e a imagem com a pele segmentada.

Figura 11 - Pele do indivíduo segmentada da imagem.

Fonte: Adaptada de Severino Junior e Gonzaga, (2006).

No trabalho de Wang, Chang e Liu (2009), é implementado um método de detecção de pele e segmentação de face humana em imagens. A primeira etapa do processo é a detecção de pele. Antes de criar um modelo de cores de pele, a imagem em RGB passa por um processo de compensação da variação de iluminação, como demonstrado na Figura 12.

Figura 12 - (a) imagem original e (b) imagem com compensação de iluminação.

Fonte: Wang, Chang e Liu (2009).

Após a compensação de iluminação, a imagem é transformada para o espaço de cores YCbCr e então é criado o modelo de cores de pele. Na parte de segmentação de pele, é aplicada uma melhora no método de Otsu, a fim de

binarizar a imagem resultante (tons de cinza) na qual os pixels próximos do valor 255 são considerados pele e os próximos de 0, não pele. O resultado final do processo é exibido na Figura 13.

Figura 13 - Resultado final, onde (a) é a região de face binarizada e (b) é a região de pele segmentada.

Fonte: Wang, Chang e Liu (2009).

A Tabela 1 exibe os valores alcançados por Wang, Chang e Liu (2009).

Tabela 1 - Resultados de Wang, Chang e Liu (2009).

Classes	Acurácia (%)
Sem balanceamento de cores	75,8
Com balanceamento de cores e sem aprimoramento	87,6
Com balanceamento de cores e com aprimoramento	94,3

Wang, Chang e Liu (2009) concluem que o método proposto é capaz de detectar faces de diferentes dimensões e que futuramente este método estará sendo aplicado em rastreamento de faces.

O trabalho publicado por Jusoh et al. (2010) apresenta um método de detecção de pele baseado em *threshold* no espaço de cores RGB e no componente H (*Hue*) do espaço de cores HSV. A Figura 14 exibe o diagrama do processo de detecção de pele.

Figura 14 - Processo de detecção de pele.

Fonte: Jusoh et al. (2010).

Os resultados do processo de segmentação de pele são apresentados apenas em imagens no trabalho de Jusoh et al. (2010), os quais podem ser observados na Figura 15.

Figura 15 - Resultados de Jusoh et al. (2010).

Fonte: Jusoh et al. (2010).

Jusoh et al. (2010) conclui que o método apresentado em seu artigo é robusto o bastante para ser aplicado em imagens com fundos (*background*) complexos e consegue detectar a pele humana com grande acurácia.

Wang, Zhang e Yao (2011) desenvolveram um método de detecção de cor de pele para ser aplicado em imagens com fundo complexo. Trata-se de uma combinação do espaço de cores RGB com o YCbCr para se detectar a área de pele e então utilizam informações da textura para remover falsos positivos.

O método foi aplicado em 500 imagens com fundo complexo, incluindo variação racial, de gênero e de iluminação. Os resultados obtidos são apresentados na Tabela 2.

Tabela 2 - Resultados de Wang, Zhang e Yao (2011).

Acurácia (%)	RGB	YCbCr	Método do Artigo
Detecção de Pele	90,3	83,7	80,1
Detecção de Fundo	71,1	82,4	88,4

Os autores concluem que o método proposto consegue remover uma grande parte dos falsos positivos da imagem.

No trabalho de Severino Junior e Gonzaga (2013) é apresentado um novo espaço de cores, o HSM, baseado no misturograma. A partir deste espaço de cores, é feita a segmentação de pele em imagens faciais e em plantas. Os resultados alcançados por este trabalho atingem 94,8% de acurácia e 6,5% de falsos positivos quando aplicados em imagens dos bancos AR e FERET.

3. MÉTODO PROPOSTO

A fim de se desenvolver um método robusto para segmentação de pele humana em imagens faciais priorizando a redução de falsos positivos, este trabalho apresenta um método que consiste das seguintes etapas:

- Aplicação de reconhecimento facial;
- Criação de banco de imagens com amostras de pele;
- Segmentação de pele em imagens utilizando RNA;
- Aplicação de redução de falsos positivos na segmentação com FPAR;
- Aplicação de redução de falsos positivos na segmentação utilizando micropadrões de texturas;

A aplicação de reconhecimento facial permite observar a importância e relevância do estudo de cores em imagens de face humana, a criação do banco de imagens possibilita o restante do trabalho de segmentação e redução de falsos positivos e as aplicações do FPAR e de textura têm o objetivo de refinar os resultados para que o objetivo do trabalho seja alcançado. Estes procedimentos são apresentados em teoria e prática a seguir.

3.1. Reconhecimento Facial

O reconhecimento facial é um termo abrangente e um grande desafio para a pesquisa, o que o torna um tema bastante estudado na área de visão computacional (JAFRI; ARABNIA, 2009).

O objetivo desta etapa do trabalho é desenvolver uma aplicação CBIR de reconhecimento facial utilizando apenas informações de cor e com isto, dar embasamento às aplicações seguintes que utilizam informações de cores em imagens faciais para a segmentação de pele humana.

Essa aplicação de CBIR foi publicada em 2012 (CASATI; RODRIGUES, 2012) e consiste em utilizar as cores quantizadas do misturograma em imagens de face, extrair características de cores e fazer comparações de imagens por meio destas características. As imagens são comparadas entre si e há um critério de similaridade que as classifica, de modo a decidir entre acerto e erro do sistema CBIR.

A aplicação de reconhecimento facial é uma comparação entre 3 diferentes abordagens. A primeira abordagem, que foi desenvolvida por Severino Junior e Gonzaga (2005), a segunda utiliza-se do esboço da imagem como descritor adicional e a terceira remove a cor quantizada em branco do vetor de características, por se tratar do fundo da imagem.

O banco de faces utilizado foi o AR (MARTINEZ; BENAVENTE, 1998). Neste banco de faces não constam pessoas de pele amarelada e de pele negra, apenas pele branca. As imagens desse banco de faces são obtidas em um ambiente altamente controlado, seguindo um padrão de fundo e distância.

Nesta aplicação são utilizadas imagens de 119 indivíduos, sendo 65 homens e 54 mulheres. De cada indivíduo são aproveitadas quatro imagens, que serão divididas em quatro classes diferentes (C1, C2, C3 e C4), como observado na Figura 16. A classe C1 é composta de uma imagem neutra, ou seja, sem expressão facial. Nas classes C2, C3 e C4 os indivíduos esboçam expressões faciais, como risos e abrir a boca.

Figura 16 - Classes de imagens do banco AR.

Fonte: Martinez e Benavente (1998).

Com as faces já separadas em classes, a comparação é feita armazenando uma das classes no banco de imagens e utilizando todas as outras como entrada, uma de cada vez. A imagem de entrada deve obter como retorno a imagem da

classe armazenada do mesmo indivíduo em questão, caso retorne uma imagem de outro indivíduo, a iteração é computada como erro.

Essa aplicação foi desenvolvida em MATLAB, na versão 7.12.0 (R2011a). O uso dessa plataforma permite o acesso ao toolbox de processamento de imagens digitais, um pacote contendo funções prontas para fins de processamento de imagens, auxiliando e agilizando o desenvolvimento de aplicações (GONZALEZ; WOODS; EDDINS, 2009).

3.1.1. Pré-Processamento

A fim de diminuir o tempo de processamento das futuras etapas do método apresentado, a primeira etapa de pré-processamento é o redimensionamento da imagem. A imagem original do banco AR possui dimensão de 768x576 pixels. A imagem é redimensionada em 25% do seu tamanho original. Este parâmetro de redimensionamento resulta na menor imagem que não altera de forma significativa os resultados. A imagem tem suas bordas removidas a fim de enquadrar melhor a face na imagem e diminuir ainda mais o seu tamanho, sem perder informação. A imagem resultante deste processo possui dimensão de 124x144 pixels. As imagens do banco de faces AR já se encontram no espaço de cores RGB, não sendo preciso nenhum tipo de transformação para a etapa de quantização das cores, uma vez que o misturograma processa imagens no espaço de cores RGB.

3.1.2. Quantização de Cores e Esboço

A quantização das cores da imagem facial visa reduzir o número de cores existentes na imagem. Uma das vantagens desta redução é a atenuação das pequenas variações de iluminação. O método de quantização de cores utilizado nesta aplicação é o misturograma. Este método foi escolhido por provar ser eficaz na quantização de cores de imagens faciais e segmentação das mesmas. Cada imagem do banco de faces AR é quantizada pelo misturograma. Como resultado desta quantização, obtém-se uma imagem de 8 cores. Na Figura 17 é exibida uma

imagem do banco de faces AR pré-processada e esta mesma imagem quantizada pelo misturograma.

Figura 17 - Imagem original pré-processada (a) e imagem quantizada (b).

Observa-se na Figura 17 a segmentação das áreas da face conforme as cores da imagem original. Pode-se observar também a redução no número de cores da imagem, que passa de mais de 16 milhões de cores para 8.

Para a segunda abordagem, que utiliza informações do esboço da face em conjunto com a imagem quantizada, é gerado um esboço considerando apenas os pixels de cor preta da camada mais significativa do misturograma. A Figura 18 exibe o esboço da face gerado pelo misturograma, a imagem da Figura 18b em questão é a mesma utilizada na Figura 18a.

Figura 18 - Imagem pré-processada (a) e esboço da imagem gerado pelo misturograma (b).

Observa-se que o esboço da face mantém muitas informações do indivíduo em questão.

3.1.3. Extração de Características

Para a primeira abordagem da aplicação, as imagens utilizadas para a extração de características são as imagens já quantizadas pelo misturograma, como exibido na Figura 17b. Nesta abordagem, a informação que é extraída das imagens é o histograma de cores. Este histograma nada mais é que a soma da quantidade de pixels de cada cor. Como as imagens quantizadas possuem oito diferentes cores, o histograma de cores terá apenas oito *bins*, um para cada cor. A Figura 19 exhibe duas imagens do banco de faces AR quantizadas pelo misturograma e seus respectivos histogramas de cores. Nota-se a divergência da quantidade de certas cores no histograma por se tratar de imagens distintas.

Este histograma de cores gera um vetor de características que pode ser utilizado para fazer as comparações entre imagens.

Figura 19 - Imagens de face quantizadas pelo misturograma e seus respectivos histogramas de cores abaixo.

Tabela 3 - Vetor de características das imagens quantizadas.

Posição	Cor	Figura 19a	Figura 19b
1	Preto	2114	1706
2	Azul	1117	944
3	Verde	1599	551
4	Ciano	1644	1027
5	Vermelho	851	1425
6	Magenta	211	1141
7	Amarelo	182	1177
8	Branco	10138	9885

Figura 20 - Histogramas das imagens acrescido dos histogramas do esboço.

Para cada cor, utiliza-se uma posição no vetor de características. Portanto, nesta primeira abordagem, os vetores de características contém 8 posições cada. A

Tabela 3 exibe os vetores de características obtidos a partir das imagens da Figura 19. Podem-se observar divergências significativas entre os dois vetores de características a partir destes valores.

Na segunda abordagem, que utiliza o histograma da imagem do esboço da face, o vetor de características passa a conter 10 posições. As oito posições da abordagem anterior (imagem quantizada) adicionando mais duas posições, a quantidade de pixels na cor branca e a quantidade de pixels na cor preta, da imagem do esboço. A imagem resume-se em apenas duas cores por se tratar de uma imagem binária.

Utilizando as mesmas imagens exibidas na Figura 19, são apresentados na Figura 20 seus respectivos histogramas acrescidos dos valores do histograma do esboço da imagem.

Na Tabela 4 são exibidos os valores do vetor de características referentes às imagens da Figura 20. Nota-se que este vetor contém 10 posições.

Tabela 4 - Vetor de características de imagens quantizadas acrescido de informações de cores do esboço da face.

Posição	Cor	Figura 20a	Figura 20b
1	Preto	2114	1706
2	Azul	1117	944
3	Verde	1599	551
4	Ciano	1644	1027
5	Vermelho	851	1425
6	Magenta	211	1141
7	Amarelo	182	1177
8	Branco	10138	9885
9	Preto do Esboço	4981	3320
10	Branco do Esboço	12875	14536

A terceira abordagem desta aplicação se utiliza apenas da quantização de cores, porém, ignorando a cor branca. Essa cor é ignorada no vetor de características por se tratar do fundo da imagem.

O vetor de características gerado a partir desta abordagem possui 7 posições, como visto na Tabela 5. Este vetor de características é referente às imagens da Figura 19, desconsiderando os valores dos pixels de cor branca.

Tabela 5 - Vetor de características de 7 posições.

Posição	Cor	Figura 19a	Figura 19b
1	Preto	2114	1706
2	Azul	1117	944
3	Verde	1599	551
4	Ciano	1644	1027
5	Vermelho	851	1425
6	Magenta	211	1141
7	Amarelo	182	1177

3.1.4. Classificação

A etapa de classificação é basicamente a comparação entre os vetores de características para todas as três abordagens. Esta comparação se dá a fim de que cada imagem apresentada como entrada na aplicação retorne a imagem da classe armazenada referente ao mesmo indivíduo.

Quando uma classe de faces é inserida no banco de imagens, os seus vetores de características são computados e salvos, para posteriormente serem comparados às imagens de faces que serão apresentadas como entrada.

A comparação ocorre de forma linear. A cada imagem de entrada apresentada, seu vetor de características também é computado. Toda imagem de entrada é comparada com todas as imagens da classe armazenada. A comparação é feita a partir do cálculo da distância euclidiana.

A imagem que resultar na menor distância entre seus vetores (maior similaridade), é tida como a resposta da entrada, ou seja, a aplicação entende que aquela imagem corresponde ao mesmo indivíduo da imagem de entrada. Se a imagem corresponder ao mesmo indivíduo é computado acerto, senão, erro da aplicação. Os resultados obtidos com as três abordagens descritas são apresentados no Capítulo 4. No item a seguir os procedimentos de criação de um banco de imagens com amostras de pele humana são apresentados.

3.2. Banco de Imagens SFA

O banco de imagens de pele facial humana SFA foi criado com o objetivo de auxiliar a pesquisa no campo de segmentação de pele humana em imagens faciais e pode ser encontrado no endereço: <http://www.sel.eesc.usp.br/sfa>.

Para a apresentação do SFA, este subcapítulo é dividido em duas partes. A primeira apresenta os procedimentos de construção do banco e a segunda apresenta os métodos que foram utilizados para validação do mesmo.

O procedimento de criação e validação deste banco de imagens apresentados nos itens 3.2.1, 3.2.2 e 3.2.3 foi publicado por Casati, Moraes e Rodrigues (2013).

Alguns exemplos das imagens e amostras do banco SFA são exibidos no Apêndice A – Exemplos de Imagens do SFA.

3.2.1. Construção do SFA

O SFA é composto por amostras de pele, amostras de não pele, as imagens originais e os *ground truths* (GT). As imagens das quais são extraídas as amostras e os GT são retiradas de bancos de imagens de face clássicos e muito utilizados na literatura: AR (MARTINEZ; BENAVENTE, 1998) e FERET (PHILLIPS et al., 2000). Exemplos destas imagens são exibidos na Figura 21.

Foi utilizado um total de 1.118 imagens, sendo 876 do banco FERET e 242 do banco AR. A distribuição destas imagens é mostrada no gráfico da Figura 22.

Figura 21 – As figuras (a) e (b) são exemplos de imagens originais do banco AR e as figuras (c) e (d) são exemplos de imagens originais do banco FERET.

Fonte: Phillips et al. (2000) e Martinez e Benavente (1998).

Figura 22 - Gráfico de distribuição da quantidade de imagens utilizadas de cada banco.

O motivo pelo qual são utilizadas mais imagens do banco FERET é que elas foram criadas em ambiente menos controlado que as imagens do banco AR, contendo variações de fundo, cor de pele, idade e sexo. Enquanto no banco AR, as variações de fundo não existem e as variações de cor de pele e idade são pouco significantes. Todas as imagens do banco SFA foram salvas em formato JPEG com qualidade 100%.

Para cada uma das 1.118 imagens, existe seu GT relativo, que foram gentilmente cedidos pelos pesquisadores Severino Jr. e Gonzaga (2013). Alguns exemplos destes GT e suas respectivas imagens originais são exibidos na Figura 23.

Figura 23 - Imagens originais seguidas de seus respectivos Ground Truths.

Fonte: adaptado de Phillips et al. (2000), Severino Jr. e Gonzaga (2013) e Martinez e Benavente (1998).

As amostras foram extraídas de cada imagem de forma aleatória, sendo 3 amostras de pele e 5 de não pele por imagem, totalizando 3.354 amostras de pele e 5.590 de não pele. Mais informações sobre a quantidade de imagens do banco SFA são exibidas na Tabela 6.

Tabela 6 - Quantidade de Imagens do SFA.

Tipo de Imagem	Quantidade
Imagens Originais	1.118
<i>Ground Truths</i>	1.118
Amostras de Pele*	3.354
Amostras de Não Pele*	5.590
Total de Imagens	163.228
*Para cada uma das 18 diferentes dimensões	

As amostras de pele e não pele contém diferentes dimensões, o que possibilita diversas investigações, como por exemplo a aplicação de textura. As dimensões das amostras variam de um pixel a 35x35 pixels, variando de dois em dois pixels.

As amostras são construídas partindo de um pixel inicial na imagem. Este pixel é pertencente à área de pele para as amostras de pele ou pertencente à área de não pele, para as outras amostras. A área a ser extraída cresce enquanto as imagens vão sendo gravadas, até que a maior amostra seja salva. Um exemplo para melhor ilustrar esta construção das amostras é exibido na Figura 24. Este procedimento é utilizado tanto para as amostras de pele quanto para as de não pele.

As imagens e as amostras contidas no SFA são organizadas em uma estrutura de pastas. São quatro principais divisões: Imagens Originais (ORI), *Ground Truths* (GT), Amostras de Pele (SKIN) e Amostras de Não Pele (NS). As pastas de amostras ainda são divididas em uma pasta para cada diferente dimensão. Esta estrutura organizacional com suas pastas e subpastas é apresentada na Figura 25.

Para enfatizar a qualidade das amostras geradas, nas próximas duas seções são apresentadas uma classificação destas amostras utilizando RNA e também a segmentação das imagens originais.

Caso haja uma aplicação que necessite de um grande número de amostras com a dimensão de apenas 1 pixel, pode-se dividir as maiores amostras (de 35x35 pixels) em amostras de apenas 1 pixel. Com isto, a quantidade de amostras de pele geradas chega a 4.108.650 e 6.847.750 de não pele.

Figura 24 - Exemplo de criação das amostras.

Figura 25 - Estrutura de pastas do SFA.

3.2.2. Validação com Classificação das Amostras

A etapa de validação das amostras é muito importante, pois demonstra o quão correlacionadas estão as amostras com suas respectivas classes (pele e não pele).

A validação é feita utilizando RNA, sistema escolhido por se tratar de uma maneira inteligente de separação e classificação do espaço amostral. A arquitetura MLP é utilizada neste processo, conhecida como a mais utilizada em processos de classificação de padrões (SILVA; SPATTI; FLAUZINO, 2010). Esta arquitetura necessita de um treinamento supervisionado para a classificação das amostras, ou seja, para cada entrada, conhece-se uma saída desejada, neste caso pele ou não pele.

As imagens do banco SFA estão em RGB, portanto, este espaço de cores é utilizado em todo o processo de validação. Sendo os valores de R, G e B as entradas para a rede neural e 0 (não pele) ou 1 (pele) a saída.

As amostras utilizadas para esta validação são as de dimensão: 1 pixel, 5x5 pixels, 15x15 pixels e 25x25 pixels. Para as amostras maiores que 1 pixel, a entrada para a rede neural é a média dos valores de R, a média dos valores de G e a média dos valores de B, mantendo a topologia da rede com três entradas.

A escolha da melhor topologia para a RNA baseia-se no método desenvolvido no trabalho de Moraes, Casati e Rodrigues (2013), variando o número de neurônios de 2 a 20 na camada escondida da rede, aumentando o seu número de 2 em 2 neurônios, como ilustrado na Figura 26.

Figura 26 - Topologias da RNA para validação das amostras.

Fonte: adaptado do *toolbox* de redes neurais artificiais do MATLAB.

Para cada uma das diferentes 10 topologias de rede, são executados 10 treinamentos e para cada treinamento, as amostras são divididas aleatoriamente em

três grupos: 70% de amostras de treinamento, 15% de amostras para validação e 15% para amostras de teste, valores padrão utilizados no *toolbox* de redes neurais do MATLAB. Os valores dos pesos e limiares de ativação são aleatoriamente definidos com valores entre 0 e 1.

Os valores R, G e B são normalizados antes de serem apresentados como entrada da rede neural, com valores entre -1 e 1. A função de ativação utilizada na camada escondida da rede é a tangente hiperbólica. A precisão requerida é definida em 10^{-4} , o número limite de épocas para o treinamento é definido em 150 e o algoritmo utilizado é o *Scaled Conjugate Gradient* (SGA) (MØLLER, 1993), que é o algoritmo padrão do *toolbox* de redes neurais do MATLAB para resolução de problemas de classificação de padrões.

Na camada de saída da rede é utilizada uma função de ativação linear e apenas um neurônio, onde, quanto mais próximo de 1 for a saída, mais próximo de pele é a amostra em questão e quanto mais próximo de 0, mais próximo de não pele é a amostra.

Para a escolha da melhor topologia, é utilizado o parâmetro de correlação entre as entradas e as saídas da rede (r), considerando apenas o grupo de amostras de teste. Outro método de validação é apresentado no próximo item, onde é feita também uma comparação do SFA com o banco UCI (BHATT; DHALL, 2013).

3.2.3. Validação com Segmentação de Imagens

Neste processo de validação, as imagens originais do banco SFA são segmentadas utilizando dois diferentes treinamentos de RNA. Um treinamento utilizando as amostras do banco SFA e outro treinamento utilizando amostras do banco UCI (“UCI *skin segmentation dataset*”) (BHATT; DHALL, 2013). O UCI é composto por 245.057 amostras RGB de pele e não pele, sendo destas 50.859 amostras de pele e 194.198 amostras de não pele. Estas amostras são extraídas de imagens dos bancos FERET (PHILLIPS et al., 2000) e PAL.

O mesmo processo de validação das amostras executado no item anterior com o SFA é aplicado também ao UCI, mas por se tratar de um banco com amostras de apenas 1 pixel, apenas a etapa com esta dimensão de amostras é executada.

Para critérios de comparação, apenas os treinamentos utilizando a dimensão de amostra de 1 pixel do SFA são considerados. O melhor treinamento para ambos os bancos de imagens é utilizado nesta etapa.

Os valores RGB das 1.118 imagens são introduzidos na RNA pixel a pixel, e a saída é um valor entre 0 e 1 para cada pixel. Esta saída da rede neural é posteriormente binarizada, resultando em uma imagem onde o valor 1 (branco) é correspondente à pele e o valor 0 (preto) à não pele. Este resultado é comparado aos GT também binarizados, como exemplificado na Figura 27 e a partir desta comparação, métricas de avaliação são obtidas.

Figura 27 - (a) Imagem original do banco SFA e (b) imagem do GT binarizada.

Fonte: Casati, Moraes e Rodrigues (2013).

A binarização da imagem resultante da RNA leva em consideração diferentes *thresholds* (cortes). Estes *thresholds* variam de 0,05% até 0,95%, valores que permitem avaliar a saída da rede neural e obter o ponto de corte ótimo para a binarização das imagens e segmentação da pele nas imagens originais.

3.3. Segmentação da Pele

Esta se trata da principal etapa do método proposto neste trabalho. Primeiramente a área com pele humana é segmentada nas imagens originais do SFA utilizando diferentes espaços de cores. Para esta etapa, foram escolhidos os espaços de cor mais utilizados em problemas de detecção de pele na literatura pesquisada (RGB, HSV, YCbCr e RGB normalizado). O espaço de cor que obtém o melhor resultado de segmentação é utilizado nas aplicações seguintes de redução de falsos positivos.

Após a segmentação das imagens por meio de uma RNA, são aplicadas técnicas de redução de falsos positivos nas imagens. Estas técnicas são denominadas FPAR (*False Positive Area Reduction*) e Textura (trata-se da análise de micropadrões de textura na imagem).

3.3.1. Segmentação por RNA em Diferentes Espaços de Cores

A segmentação das imagens de faces originalmente no espaço de cores RGB podem sofrer alterações quando este espaço de cores é convertido a outros. Alguns trabalhos abordam a comparação dos espaços de cores em segmentação de pele humana em imagens (CHAVES-GONZÁLEZ et al., 2010; KAKUMANU; MAKROGIANNIS; BOURBAKIS, 2007), porém, para que a aplicação seja direcionada ao uso do SFA, são executados testes de segmentação nos espaços de cores mais utilizados na literatura.

As imagens estão originalmente no espaço de cores RGB, que é o espaço de cores padrão dos bancos de imagens AR e FERET. É necessária a conversão das imagens e das amostras para os espaços HSV, YCbCr e RGB Normalizado (NormRGB) para que se possa avaliar a diferença na segmentação das imagens e posteriormente escolher o melhor resultado para que se possa aplicar outras técnicas de refinamento.

A segmentação das imagens é feita por meio de uma RNA treinada com as amostras de 1 pixel, seguindo os mesmos procedimentos utilizados nas seções 3.2.2 e 3.2.3, baseados no trabalho de Moraes, Casati e Rodrigues (2013), com a única diferença que nesta etapa são utilizadas duas camadas escondidas na RNA,

variando de 1 a 20 neurônios cada, visando um maior número de treinamentos até que se encontre o treinamento ótimo para a segmentação. A topologia é ilustrada na Figura 28.

Figura 28 - Topologias de RNA utilizadas para segmentação de pele nas imagens.

Fonte: adaptado do *toolbox* de redes neurais artificiais do MATLAB.

A cada uma das 400 diferentes topologias testadas, são executados 10 treinamentos, totalizando 4.000 treinamentos, para cada um dos espaços de cores. Para todos os espaços de cores testados, o número de entradas não se altera, mantendo-se sempre três entradas para a rede neural:

- Entradas RGB: (*Red*, *Green*, *Blue*);
- Entradas HSV: (*Hue*, *Saturation*, *Value*);
- Entradas YCbCr: (*Y*, *Cb*, *Cr*);
- Entradas normRGB: (*normRed*, *normGreen*, *normBlue*).

O melhor treinamento da melhor topologia é escolhido de acordo com (MORAES; CASATI; RODRIGUES, 2013), ou seja, levando em consideração o $r_{\max}$.

O resultado deste processo (saída da RNA) é uma imagem em escala de cinza normalizada entre 0 e 1, onde quanto mais próximo de 1 for o valor do pixel, mais se assemelha com pele e quanto mais próximo de 0, mais se assemelha com não pele. Nesta imagem de saída são efetuados diversos cortes (*thresholds*) até que se encontre um *threshold* ótimo, ou seja, que apresenta o melhor resultado quando comparado aos GT. Este *threshold* varia de 0,05 a 0,95 (ou seja, de 5% a 95%) incrementando em 0,05 (5%) por iteração.

O *threshold* ótimo é dado pelo valor que obteve a maior acurácia média segmentação das 1.118 imagens testadas. A imagem resultante deste procedimento é utilizada como entrada para as outras aplicações deste trabalho (aplicação de FPAR e Textura), apresentadas a seguir no item 3.3.2.

3.3.2. Aplicação do FPAR

A aplicação do FPAR tem como objetivo reduzir os falsos positivos das imagens já segmentadas sem que isto diminua a acurácia alcançada na etapa anterior.

O resultado da aplicação apresentada no item 3.3.1, considerando apenas espaço de cores que obteve os melhores resultados, é a entrada para este procedimento: a imagem binarizada pelo melhor *threshold*, onde o branco corresponde à pele e o preto a não pele. Um diagrama para melhor exemplificar o procedimento do FPAR é exibido na Figura 29.

Figura 29 - Diagrama de processos do FPAR.

O procedimento do FPAR consiste em reduzir os segmentos detectados como pele que não são contínuos, ou seja, que tenham o número de pixels 8-conectados menor que a porcentagem total de pixels detectados como pele. Esta porcentagem inicia de 5% e vai até 95%, variando de 5 em 5 por cento.

Este procedimento é aplicado na imagem já binarizada visando à redução de ruídos que possam ter sido detectados como pele pela rede neural, por conterem pixels com valores próximos aos valores das amostras de pele. Isto se dá porque as regiões de pele em imagens normalmente são grandes áreas contínuas (SIQUEIRA; SCHWARTZ; PEDRINI, 2013).

Exemplo de aplicação do FPAR (procedimento executado para cada imagem):

- Passo 1: A entrada é a imagem resultado da segmentação advinda da RNA, ou seja, uma imagem binária onde os pixels com valor 1 são os classificados como pele e os pixels com valor 0 são os classificados como não pele. Então, conta-se o número de pixels classificados como pele da referida imagem, ou seja, o total de positivos ($VP + FP = TP$).
- Passo 2: Calcula-se a porcentagem do TP, varrendo de 5% até 95% variando de 5 em 5 por cento.
- Passo 3: Aplica-se, para cada uma das porcentagens, a remoção das áreas com número menor de pixels aos do valor da porcentagem e que não estejam 8-conectados. Ou seja, supondo que o processo estivesse removendo as áreas menores que 10% do total de pixels classificados como pele, as áreas que serão desclassificadas, ou seja, passarão a valer 0 (não pele), serão aqueles blocos contínuos menores que 10% do TP, e assim por diante. Isto significa que nenhuma área contínua classificada como pele seria menor que 10% do TP, neste exemplo.
- Passo 4 (grupo de imagens): Para a análise do grupo de 1.118 imagens, opta-se pelo valor de porcentagem que obteve os melhores resultados levando em consideração a acurácia média entre todas as imagens.

Os resultados obtidos para o grupo de 1.118 imagens e alguns exemplos de redução dos valores de falso positivo são exibidos no Capítulo 4.

A seguir, a utilização de textura para a redução de falsos positivos é utilizada e sua aplicação é feita tanto nas imagens resultantes da RNA utilizadas como entrada neste processo quanto nas imagens resultantes do FPAR.

3.3.3. Aplicação de Textura

Nesta abordagem do trabalho, a aplicação de textura nas imagens é uma maneira de se buscar a redução de falsos positivos nos processos de detecção de pele, visto que este é um dos maiores problemas encontrados nestes sistemas. Isto acontece devido à ocorrência de pixels que não fazem parte do segmento de pele em imagens possuir cor de pele, ou seja, valores muito próximos ou iguais aos

valores de pixels de pele, induzindo assim a uma classificação errônea (SIQUEIRA; SCHWARTZ; PEDRINI, 2013).

Neste caso, procurou-se analisar a textura de uma pequena área da imagem em comparação com texturas de pele e calcular um valor de distância entre as duas. Este valor representa o quão próximo aquela textura é de uma textura de pele e posteriormente, áreas que possuem valores de textura muito distantes dos valores de textura encontrados nas amostras de pele, passam a ser classificados como não pele.

As dimensões de amostras utilizadas neste processo são: 5x5 pixels, 11x11 pixels, 15x15 pixels, 25x25 pixels e 35x35 pixels. Para esta abordagem, a imagem resultante da RNA é denominada IRNA e a imagem resultante do procedimento de redução de falsos positivos (FPAR) é denominada IFPAR.

Na análise dos micropadrões de textura é utilizado o LFP (*Local Fuzzy Pattern*). Este método foi desenvolvido por Vieira (2013), que obteve resultados promissores na descrição de texturas. A análise utilizando LBP foi descartada por não apresentar variação quando analisadas as amostras de pele e não pele, portanto, tornou-se inviável sua aplicação para esta distinção. Na Figura 30 é exibido o diagrama do processo de análise da textura.

Figura 30 - Diagrama dos processos para análise de textura.

Todas as imagens e amostras a serem trabalhadas pelo LFP são anteriormente transformadas para escala de cinza. A função de pertinência utilizada é a sigmoide, e a janela de vizinhança analisada é de 3x3 pixels.

O primeiro passo para a aplicação de textura é aplicar o LFP em todas as amostras de pele da dimensão em questão. Para cada amostra, um histograma é gerado, descrevendo a textura. Posteriormente, é calculada a média dos histogramas, o que resulta no histograma médio das amostras de pele (HMP). Cada diferente dimensão de amostras analisadas possui seu HMP, que é utilizado posteriormente para que se possa gerar a imagem de distância chi-quadrado (PELE; WERMAN, 2010).

Após a aquisição dos valores do HMP, é aplicado o LFP na imagem original, um exemplo do resultado desta aplicação é exibido na Figura 31.

Figura 31 - Imagem original do banco SFA (a) e imagem resultante da aplicação do LFP (b).

Por se tratar de um método parametrizável e por não ter até o momento uma técnica que auxilie na definição de seu parâmetro β , o processo todo é executado de forma empírica. O β varia de 0,05 a 2 com passos de 0,05 totalizando 40 valores diferentes.

O próximo passo é percorrer a imagem resultante do LFP com uma máscara de mesma dimensão da amostra que está sendo testada e, a cada iteração são efetuados os seguintes processos:

- Cálculo do histograma local, ou seja, o histograma da área coberta pela máscara no momento;
- Comparação do histograma local com o HMP para a referida dimensão de amostra, utilizando a distância chi-quadrado;
- Aplicar o valor resultante da distância chi-quadrado no pixel central da máscara.

Após estes procedimentos serem executados em toda a imagem, têm-se como resultado uma matriz com os valores da distância chi-quadrado entre a imagem original e o HMP, calculada conforme a equação 8.

$$\chi^2(P_i, Q_i) = \frac{1}{2} \sum_{i=0}^{255} \frac{(P_i - Q_i)^2}{(P_i + Q_i)} \quad (8)$$

onde Q_i é a frequência do nível de cinza referente ao histograma da área coberta pela janela a varrer a imagem e P_i é a frequência do nível de cinza do HMP.

A matriz (imagem) de distância chi-quadrado gerada é normalizada com valores entre 0 e 1, sendo que os valores mais próximos a 0 (preto) representam que os pixels em questão estão contidos em uma área com a textura mais próxima à de pele (menor distância) e os valores mais próximos a 1 (branco) representam que os pixels em questão estão contidos em uma área com a textura mais distante à textura de pele (maior distância). A Figura 32 exibe um exemplo de imagem original do banco SFA seguida de suas respectivas imagens de distância chi-quadrado (IMG_DIST) para todos os tamanhos de amostra testados neste trabalho.

Para o desenvolvimento das próximas etapas, deve ser calculado o complemento das imagens de distância, para que quando multiplicada a máscara de textura pelas imagens resultantes IRNA e IFPAR as áreas com textura mais distantes da textura de pele sejam descartadas.

As imagens de distância são binarizadas obedecendo a um *threshold* que varia de 5% a 95% do seu valor, ou seja, de 0,05 a 0,95. Após esta binarização a imagem resultante é considerada uma máscara de textura (TEX), que posteriormente é multiplicada pelas imagens IRNA e IFPAR.

Para cada diferente *threshold*, a máscara TEX é multiplicada por IRNA e também por IFPAR, resultando em duas diferentes imagens finais. Após este processo, as imagens são comparadas aos seus respectivos GT e desta comparação são extraídos os resultados do processo, apresentados no próximo capítulo.

Figura 32 - (a) Imagem original em escala de cinza, (b) IMG_DIST com amostras de 5x5, (c) IMG_DIST com amostras de 11x11, (d) IMG_DIST com amostras de 15x15, (e) IMG_DIST com amostras de 25x25, (f) IMG_DIST com amostras de 35x35.

4. RESULTADOS

Neste capítulo são apresentados os resultados dos procedimentos executados no capítulo anterior. Primeiramente, os resultados para a aplicação de reconhecimento facial por meio de cores quantizadas. Posteriormente, são apresentados os resultados da validação do banco de imagens SFA e, finalmente, são apresentados os resultados dos procedimentos de segmentação das imagens faciais com e sem a aplicação das técnicas de redução de falsos positivos.

4.1. Reconhecimento Facial

Os resultados obtidos a partir desta aplicação são apresentados por meio de tabelas. Cada tabela de resultados é referente a uma abordagem diferente. Para cada abordagem, as 4 classes de imagens foram armazenadas, e para cada armazenamento, todas as outras 4 classes foram utilizadas como entrada, em uma abordagem de comparação todas contra todas.

Para a primeira abordagem, que possui o vetor de características de 8 posições, uma para cada cor quantizada pelo misturograma, os resultados obtidos são apresentados na Tabela 7. Observa-se que quando uma classe é armazenada e a mesma classe é a entrada, os resultados sempre são de 100% de acertos, pois a distância dos vetores de características da mesma imagem é zero, portanto, são ignorados na análise dos resultados.

A primeira linha da tabela exhibe as classes armazenadas no banco de imagens e a primeira coluna as classes de faces utilizadas como entrada. Os resultados aferidos para esta primeira abordagem variam de 89,1% a 97,5%. Portanto, sofrem uma variação máxima de 8,4%. A média de acertos desta abordagem é de 92,3%.

Tabela 7 - Acertos da abordagem com vetor de 8 posições.

Classes	C1	C2	C3	C4
C1	100,0 %	97,5 %	92,4 %	92,4 %
C2	97,5 %	100,0 %	90,8 %	89,1 %
C3	92,4 %	90,8 %	100,0 %	89,9 %
C4	92,4 %	89,1 %	89,9 %	100,0 %

Na segunda abordagem, que utiliza além das 8 cores quantizadas pelo misturograma, o histograma do esboço binário da face, os resultados obtidos são exibidos na Tabela 8.

Tabela 8 - Acertos da abordagem com vetor de 10 posições.

Classes	C1	C2	C3	C4
C1	100,0 %	96,6 %	85,7 %	84,9 %
C2	96,6 %	100,0 %	89,1 %	85,7 %
C3	85,7 %	89,1 %	100,0 %	86,5 %
C4	84,9 %	85,7 %	86,5 %	100,0 %

Nesta segunda abordagem, a média de acertos é de 88,1% e os resultados variam de 96,6% a 84,9%. Com isso, a variação máxima dos resultados foi de 11,7%.

A terceira abordagem, que utiliza apenas 7 das 8 cores quantizadas pelo misturograma, ignorando a cor branca por se tratar do fundo da imagem no banco de faces utilizado, tem seus resultados apresentados na Tabela 9.

Esta abordagem alcança média de acertos de 92,7% e sua variação máxima dos resultados é de 6,7%.

Os resultados apresentados permitem avaliar positivamente a abordagem que utiliza apenas 7 posições no histograma de cores, pois na média, foi a abordagem que obteve os melhores resultados.

Tabela 9 - Acertos da abordagem com vetor de 7 posições.

Classes	C1	C2	C3	C4
C1	100,0 %	96,6 %	92,4 %	92,4 %
C2	96,6 %	100,0 %	91,6 %	93,3 %
C3	92,4 %	91,6 %	100,0 %	89,9 %
C4	92,4 %	93,3 %	89,9 %	100,0 %

A respeito do desempenho quanto ao custo computacional das três abordagens, os resultados obtidos foram aferidos em um notebook com as características a seguir: processador *Core i7 Q720 @ 1.6 GHz*, 4 GB de memória RAM DDR3 e sistema operacional *Windows 7* de 64 *bits*, com a aplicação rodando em ambiente convencional.

A Tabela 10 apresenta os valores em segundos do tempo de processamento para cada etapa, das três abordagens desenvolvidas. A medida é uma média do processamento para cada imagem. A abordagem com vetor de 8 posições é denominada A8, a abordagem com vetor de 10 posições é denominada A10 e a que utiliza vetor de 7 posições é chamada de A7.

Tabela 10 - Medidas de velocidade de processamento em segundos.

Processo	A8	A10	A7
Pré-processamento e quantização de cores	3,62	3,68	3,62
Classificação	0,06	0,06	0,06

Como se pode observar, não houveram diferenças significantes no tempo de processamento independente do tamanho do vetor utilizado, porém, além de obter os melhores resultados, a abordagem com 7 posições possui menos dados para serem processados.

O objetivo da aplicação é alcançado, atingindo altas taxas de acertos para reconhecimento facial utilizando apenas cores como descritores, ratificando a tese de que as cores são altamente descritivas em imagens de faces humanas.

4.2. Banco de Imagens SFA

A validação das amostras faz-se necessária para que se possa observar o comportamento das amostras em uma aplicação real e mensurar o seu desempenho quanto à sua aplicação em segmentação de pele humana em imagens. Primeiramente, são apresentados os resultados da classificação das amostras, onde pode-se obter uma visão da correlação das mesmas com suas respectivas classes. Posteriormente, os resultados da aplicação em segmentação de imagens são exibidos, incluindo uma comparação com o desempenho de outro banco de amostras de pele.

4.2.1. Validação com Classificação das Amostras

Os resultados desta etapa de validação são muito importantes, pois por meio deles podem-se analisar as amostras de forma separada, ou seja, o quanto estas amostras representam a realidade do universo do qual foram extraídas.

O melhor treinamento das amostras foi escolhido com base no maior valor de correlação ($r_{\max}$) entre as amostras e suas respectivas classes, considerando apenas o grupo de amostras de teste (15% do total de amostras).

A Tabela 11 mostra os valores $r_{\max}$ para cada uma das diferentes dimensões utilizadas neste procedimento, considerando o melhor treinamento. A topologia deste treinamento também é exibida na Tabela 11, considerando apenas o número de neurônios na camada escondida, que é a única variação entre as diferentes topologias testadas.

Tabela 11 - Resultados da validação para grupo de amostras de teste da RNA.

Dimensão	Neurônios	Correlação ($r_{\max}$)	Acurácia
1 pixel	20	0,9333	96,02%
5x5 pixels	20	0,9312	96,00%
15x15 pixels	20	0,9312	95,96%
25x25 pixels	10	0,9206	95,54%

A acurácia apresentada na tabela dá-se pelo total de amostras do grupo de testes que foram corretamente classificadas.

Os resultados apresentam altos valores de correlação e acurácia em todos os casos, mostrando que as amostras geradas para o banco de imagens SFA têm alta correlação com suas classes, não havendo variação significativa entre as diferentes dimensões.

Nota-se também uma ligeira queda no desempenho ao aumentar a dimensão das amostras, porém, a maior dimensão testada (25x25 pixels) apresenta resultados muito próximos das outras dimensões com metade dos neurônios em sua camada escondida, diminuindo assim o custo computacional do processo.

4.2.2. Validação com Segmentação de Imagens

Os resultados são apresentados levando em consideração o melhor treinamento da melhor topologia para cada um dos bancos (UCI e SFA).

A imagem segmentada é comparada aos GT e extraída a acurácia para cada procedimento, ou seja, o percentual de pixels classificados corretamente como pele ou não pele. A Tabela 12 exibe os resultados para ambos os bancos, incluindo também o *threshold* utilizado para cada um, no caso, o que ofereceu o melhor resultado.

Tabela 12 - Resultados de acurácia para a segmentação utilizando SFA e UCI.

Base de Imagens	Melhor <i>Threshold</i>	Acurácia
SFA	10%	92,71%
UCI	10%	88,74%

Observa-se que o melhor *threshold* para ambos os bancos foi de 0,10, ou seja, o corte é feito em 10% do valor máximo possível que o pixel pode ter (no caso, o valor de 255).

Para um melhor entendimento do comportamento do *threshold* aplicado à imagem resultante da RNA, é apresentado um gráfico na Figura 33 onde os resultados de acurácia são exibidos para todos os *thresholds* aplicados.

Figura 33 - Comportamento da acurácia em relação ao *threshold* aplicado para ambos os bancos (UCI e SFA).

Pode-se perceber analisando a Tabela 12 e a Figura 33 que as amostras do banco SFA superaram as amostras do banco UCI em torno de 4% com *threshold* de 0,10, o que mostra que as amostras do banco SFA aplicadas no treinamento da rede neural foram mais eficazes que as amostras do banco UCI, para as 1.118 imagens analisadas.

Na Figura 34 são exibidos alguns exemplos de imagens segmentadas pelas amostras dos dois bancos testados.

Ao observar a Figura 34, nota-se a diferença no desempenho entre a segmentação pelas amostras do UCI e do SFA, enfatizando ainda mais que as amostras do banco SFA são eficazes na segmentação das imagens testadas.

Figura 34 - Comparação entre a segmentação utilizando SFA e UCI. A primeira coluna (a) contém as imagens originais, a segunda coluna (b) contém os GT, a terceira coluna (c) contém as imagens segmentadas utilizando amostras do SFA e a quarta coluna (d) contém as imagens segmentadas utilizando as amostras do UCI.

4.3. Segmentação da Pele

Neste item são exibidos os resultados das aplicações de segmentação de pele humana em imagens. Estes resultados são divididos em três partes, assim como a divisão da explicação do método proposto. Na primeira (4.3.1), são exibidos

os resultados da segmentação das 1.118 imagens originais do SFA em quatro diferentes espaços de cores. No item 4.3.2 são apresentados resultados da aplicação do FPAR e no item 4.3.3, os resultados da aplicação de textura nas imagens resultantes da RNA (IRNA) e nas imagens resultantes da aplicação de FPAR (IFPAR).

4.3.1. Segmentação por RNA em Diferentes Espaços de Cores

Os testes foram realizados com quatro diferentes espaços de cores e seu desempenho é medido levando em consideração a acurácia média das 1.118 imagens originais segmentadas.

Na Tabela 13 pode-se analisar a topologia da RNA que ofereceu a maior correlação para cada um dos quatro espaços de cores testados.

Tabela 13 - Valores de correlação por espaços de cores e topologia da RNA.

Espaço de Cores	Correlação ($r_{\max}$)	Topologia da RNA
		(Primeira camada escondida / segunda camada escondida)
RGB	0,9349	17 / 19
HSV	0,9254	14 / 18
YCbCr	0,9333	10 / 7
NormRGB	0,8486	13 / 19

Nota-se que não há diferença significativa nos valores de correlação de RGB, HSV e YCbCr, porém, o NormRGB não foi capaz de produzir um valor de correlação próximo ao dos outros espaços de cores.

A Tabela 14 exibe os resultados de acurácia da segmentação para os quatro espaços de cores, considerando o banco SFA completo (SFA), apenas as 876 imagens do banco FERET (IM_FERET), apenas as 242 imagens do banco AR (IM_AR) e a média da acurácia entre os valores de IM_FERET e IM_AR. Estes resultados são utilizados para a definição do melhor espaço de cores para este processo e o mesmo será utilizado nos demais procedimentos do trabalho.

Tabela 14 - Acurácia para segmentação das imagens por espaço de cores.

Espaço de Cores	SFA	IM_FERET	IM_AR	Média
RGB	92,74%	91,59%	96,92%	94,26%
HSV	62,10%	66,07%	47,67%	56,87%
YCbCr	75,76%	72,27%	88,42%	80,35%
NormRGB	73,02%	72,35%	75,43%	73,89%

Pode-se observar na Tabela 14 que a menor acurácia é alcançada pelo espaço de cores HSV. Mesmo tendo na literatura trabalhos onde o espaço de cores HSV é defendido como melhor que o RGB para segmentação de pele (SIQUEIRA; SCHWARTZ; PEDRINI, 2013), neste caso a acurácia alcançada por este espaço de cores é significativamente menor que a alcançada pelo RGB. Isto se dá devido ao fato de as imagens terem sido convertidas para JPEG que, mesmo mantendo esta compressão com 100% da qualidade, alguns pixels têm seus valores alterados e foi observado neste trabalho que isto ocorreu com os pixels brancos. Por este motivo, estes pixels brancos foram classificados como pele. Este processo inviabiliza o uso do HSV para as imagens testadas neste trabalho.

Os espaços de cores YCbCr e o NormRGB não alcançaram valores próximos ao RGB e portanto foram também descartados.

Observa-se uma maior acurácia para o espaço de cores RGB para todos os tipos de separação do banco de imagens. As imagens do banco FERET possuem menor acurácia, isto se dá devido à maior variação de ambientes e indivíduos nas imagens. Já no banco AR, as imagens são mais bem comportadas, possuindo o mesmo fundo e a cor da pele dos indivíduos varia menos.

Observa-se no gráfico da Figura 35 que o espaço de cores RGB tem um desempenho consideravelmente maior que os outros espaços de cores.

É importante acrescentar também que o percentual de falsos positivos (FP) desta segmentação por RNA para o espaço de cores RGB é um pouco elevado, como observado na Tabela 15.

Figura 35 - Gráfico de acurácia dos espaços de cores.

Tabela 15 - Resultados de FP para a aplicação de RNA no espaço de cores RGB.

SFA	11,42%
IM_FERET	11,59%
IM_AR	10,81%
Média	11,20%

Resultados considerando as aplicações que visam à redução destes índices de falsos positivos são exibidos e discutidos a seguir.

4.3.2. Aplicação do FPAR

A partir deste ponto são apresentados os resultados considerando a acurácia dos métodos e também a quantidade de falsos positivos. Esta quantidade tende a diminuir com a aplicação do FPAR, pois há remoção de áreas não contínuas na imagem que normalmente não fazem parte do segmento de pele humana em imagens faciais.

Para cada um dos procedimentos de remoção, ou seja, para cada percentual de VP que é removido da imagem como explicado no item 3.3.2, são exibidos os valores de acurácia na Tabela 16 e os valores de falsos positivos (FP) na Tabela 17.

Tabela 16 - Valores de acurácia com aplicação do FPAR.

Quantidade de VP	SFA	IM_FERET	IM_AR	Média
5%	93,28%	92,04%	97,76%	94,90%
10%	93,37%	92,11%	97,93%	95,02%
<u>15%</u>	<u>93,42%</u>	<u>92,14%</u>	<u>98,03%</u>	<u>95,09%</u>
20%	93,37%	92,09%	98,01%	95,04%
25%	93,27%	91,98%	97,96%	94,97%
30%	93,25%	91,96%	97,94%	94,95%
35%	93,22%	91,91%	97,94%	94,93%
40%	93,21%	91,90%	97,94%	94,92%
45%	93,19%	91,88%	97,94%	94,91%
50%	93,18%	91,86%	97,94%	94,90%
55%	93,15%	91,82%	97,94%	94,88%
60%	93,10%	91,76%	97,94%	94,85%
65%	92,97%	91,62%	97,84%	94,73%
70%	92,70%	91,31%	97,74%	94,53%
75%	92,19%	90,80%	97,25%	94,02%
80%	91,71%	90,28%	96,89%	93,59%
85%	90,88%	89,49%	95,90%	92,69%
90%	89,67%	88,39%	94,29%	91,34%
95%	87,01%	85,84%	91,22%	88,53%

Na Tabela 16, a primeira coluna se refere à porcentagem que é calculada do total de VP. O resultado deste cálculo é a quantidade mínima de pixels que os blocos 8-conectados de VP, ou seja, os blocos classificados como pele, devem ter. Em outras palavras, os blocos classificados como pele que forem menores que este valor são reclassificados como não pele. A Figura 36 exibe o comportamento referente à Tabela 16 para cada um dos grupos de imagens.

Pode-se observar que a acurácia tem poucas variações e só começa a diminuir consideravelmente após a remoção dos blocos menores que 60% do total de VP da imagem segmentada.

Figura 36 - Gráfico do comportamento da acurácia com a aplicação do FPAR.

O pico máximo de acurácia se dá no ponto onde são removidos blocos menores que 15% (destacado na Tabela 16) do total de VP na imagem, ou seja, este é o valor ótimo alcançado pelo FPAR quando se considera a acurácia da segmentação das imagens. O mesmo valor ótimo é encontrado tanto para todo o banco de imagens SFA quanto para as imagens IM_AR e IM_FERET.

Na Tabela 17 são exibidos os valores de FP para cada uma das iterações. Pode-se observar que na medida em que o tamanho dos blocos a serem removidos aumenta, o número de FP diminui. Porém, paga-se um preço com a diminuição da acurácia, em outras palavras, a quantidade de VP também é reduzida. Nota-se também uma queda mais acentuada entre os valores de 5% e 15%, após isto, a diminuição dos FP não é tão efetiva.

O gráfico da Figura 37 exhibe o comportamento do FPAR segundo os valores de FP, onde se pode observar a diminuição gradativa destes valores. Observa-se também que a redução de FP é mais acentuada até o valor de 15%.

A aplicação do FPAR mostrou-se eficiente tanto para o aumento da acurácia quanto para a redução de falsos positivos, melhorando os resultados nas duas métricas.

Tabela 17 - Resultados de FP para a aplicação do FPAR.

Quantidade de VP	SFA	IM_FERET	IM_AR	Média
5%	9,09%	9,67%	7,01%	8,34%
10%	8,54%	9,21%	6,14%	7,68%
15%	8,03%	8,77%	5,34%	7,06%
20%	7,87%	8,60%	5,22%	6,91%
25%	7,82%	8,54%	5,22%	6,88%
30%	7,65%	8,32%	5,22%	6,77%
35%	7,62%	8,29%	5,22%	6,76%
40%	7,62%	8,28%	5,22%	6,75%
45%	7,62%	8,28%	5,22%	6,75%
50%	7,62%	8,28%	5,22%	6,75%
55%	7,62%	8,28%	5,22%	6,75%
60%	7,53%	8,16%	5,22%	6,69%
65%	7,43%	8,04%	5,21%	6,63%
70%	7,36%	7,96%	5,18%	6,57%
75%	7,13%	7,73%	4,98%	6,36%
80%	6,92%	7,53%	4,69%	6,11%
85%	6,64%	7,27%	4,38%	5,83%
90%	6,17%	6,73%	4,12%	5,43%
95%	5,11%	5,65%	3,15%	4,40%

Para a aplicação de textura, foram escolhidas as imagens resultantes da aplicação do FPAR com remoção dos blocos menores que 15% do total de VP de cada imagem. Esta escolha é feita para manter a maior acurácia alcançada para posteriormente aplicar a redução de FP utilizando textura. Este parâmetro também é utilizado na comparação com a aplicação de RNA que é feita no gráfico da Figura 38.

Figura 37 - Gráfico do comportamento dos FP com a aplicação do FPAR e parâmetro de 15%.

Figura 38 - Comparativo de FP entre RNA e FPAR.

A redução de falsos positivos é de 29,7% para todas as imagens (SFA) utilizando o parâmetro que obteve a melhor acurácia, ou seja, sem que a precisão do método seja prejudicada, pelo contrário, esta precisão aumentou.

No próximo item são exibidos os resultados finais deste trabalho com comparativos entre as aplicações desenvolvidas.

4.3.3. Aplicação de Textura

Este item apresenta os resultados do processo de aplicação de textura nas imagens do banco SFA. O procedimento é executado nas imagens resultantes da RNA e nas imagens resultantes da aplicação do FPAR. É apresentado também um comparativo dos resultados dos diferentes procedimentos que são parte do método desenvolvido neste trabalho.

Nesta etapa do método, a aplicação de textura, visa-se reduzir a quantidade de falsos positivos sem que a acurácia tenha diminuição expressiva. Para tanto, é aplicada a análise de micropadrões nas imagens do SFA, utilizando LFP e os resultados alcançados com esta abordagem são exibidos neste item.

Na Tabela 18 são apresentados os resultados da média da acurácia e da média dos falsos positivos considerando os 40 valores de β testados, após a aplicação de textura nas imagens resultantes da RNA (IRNA).

Ao analisá-la, vê-se que não há grande variação entre os resultados para cada dimensão de amostra, porém, à medida que a dimensão da amostra aumenta, a quantidade de FP é reduzida. Por outro lado, a acurácia também é reduzida.

Tabela 18 - Resultados da aplicação de textura nas imagens IRNA.

	5x5	11x11	15x15	25x25	35x35
Acurácia	92,06%	92,52%	92,50%	92,37%	91,80%
Falsos Positivos	8,06%	8,11%	7,98%	7,73%	7,39%

O gráfico apresentado na Figura 39 exibe os valores de acurácia alcançados para cada valor de β testado, resultantes da aplicação de textura nas imagens IRNA, para as cinco diferentes dimensões de amostras.

A divergência significativa entre a acurácia de cada dimensão de amostra começa a aparecer após $\beta = 0,25$, acarretando em uma diminuição da acurácia mais acentuada para dimensões de amostra: 5x5 e 35x35.

Figura 39 - Acurácia em relação aos valores de β para textura aplicada em IRNA.

Na Figura 40 é apresentado o gráfico do comportamento dos FP na variação dos valores de β testados.

Figura 40 – Falsos positivos em relação aos valores de β para textura aplicada em IRNA.

Os valores de FP diminuem significativamente após $\beta = 0,25$, acompanhando o comportamento do β quanto à acurácia do processo.

Ao analisar os gráficos pode-se determinar que os valores de β reproduzem resultados muito próximos de $\beta = 0,05$ a $\beta = 0,2$ e começam a variar significativamente após o $\beta = 0,25$. Observa-se também que, embora a variação não seja brusca, ela é significativa, o que permite concluir que o parâmetro β é relevante para o processo.

A Tabela 19 exibe os resultados ótimos alcançados pela aplicação de textura nas imagens IRNA e seus respectivos parâmetros, considerando o maior valor de acurácia e também o menor valor de FP.

Tabela 19 - Resultados ótimos da aplicação de textura nas imagens IRNA.

Abordagem	β	Threshold	Amostra	Acurácia	FP
Maior Acurácia	2,00	95%	11x11	92,54%	8,08%
Menor FP	1,65	95%	35x35	91,68%	7,06%

Os resultados ótimos exibidos na Tabela 19 permitem escolher os parâmetros indicados para os objetivos da aplicação, seja priorizando a redução de falsos positivos ou esta redução mantendo os níveis de acurácia um pouco mais elevados.

A aplicação de textura também ocorre nas imagens resultantes da aplicação do FPAR (IFPAR). Os resultados desta abordagem encontram-se na Tabela 20. Estes resultados tratam-se da média de todos os valores de β aplicados às imagens, para cada diferente dimensão de amostra.

Tabela 20 - Resultados da aplicação de textura nas imagens IFPAR.

	5x5	11x11	15x15	25x25	35x35
Acurácia	92,98%	93,39%	93,30%	92,91%	91,88%
Falsos Positivos	7,72%	7,54%	7,34%	6,80%	6,19%

Os resultados superam os da aplicação de textura nas imagens IRNA. Isto é esperado, pois as imagens resultantes do FPAR com o parâmetro de remoção em 15% obtêm maior acurácia e menor percentual de FP que as IRNA, ou seja, superam nas duas métricas consideradas neste trabalho.

O comportamento do parâmetro β em relação à acurácia, aplicado às imagens IFPAR pode ser observado no gráfico da Figura 41.

Figura 41 - Acurácia em relação aos valores de β para textura aplicada em IFPAR.

Observa-se um comportamento muito parecido com o demonstrado na Figura 39, relativo à aplicação de textura nas imagens IRNA, com exceção das amostras de dimensão 35x35 que alcançam uma acurácia significativamente inferior às demais. Porém, analisando-se os resultados, fica claro que a acurácia alcançada quando da aplicação de textura nas imagens IFPAR é superior à acurácia alcançada com as imagens IRNA.

Figura 42 - Falsos positivos em relação aos valores de β para textura aplicada em IFPAR.

O mesmo comportamento da acurácia, que obteve melhores resultados com a textura aplicada às IFPAR, também é esperado para os FP, visto que as imagens IFPAR contém percentual menor destes. A Tabela 20, a Tabela 21 e a Figura 42 ratificam este comportamento, mostrando que a aplicação de textura nas imagens IFPAR diminuiu o percentual de FP, atingindo valores mais baixos que o resultado da aplicação de textura nas imagens IRNA.

Os valores de FP não sofrem alteração significativa com a variação do β , como pode-se observar na Figura 42. É possível observar também que quanto maior a dimensão da amostra utilizada, mais a porcentagem de FP é reduzida. Comportamento semelhante foi verificado na análise da acurácia (Figura 41).

Os resultados ótimos e seus parâmetros para a aplicação de textura nas IFPAR são apresentados na Tabela 21, considerando a maior acurácia alcançada e também o menor percentual de FP alcançado.

Tabela 21 - Resultados ótimos da aplicação de textura nas imagens IFPAR.

Abordagem	β	Threshold	Amostra	Acurácia	FP
Maior Acurácia	2,00	90%	11x11	93,42%	7,49%
Menor FP	1,65	95%	35x35	91,69%	6,06%

O menor percentual de FP atingido com o método proposto neste trabalho é com a aplicação de textura nas imagens FPAR (com exceção da aplicação de FPAR e parâmetro de remoção em 95%, porém, a acurácia desse processo é de 87%, significativamente menor que a alcançada com a aplicação de textura nas IFPAR). Isto fica claro analisando a última linha da Tabela 21, onde o percentual de falsos positivos alcança 6,06%. Este valor significa 43,9% de redução em relação à segmentação por RNA e 24,5% de redução em relação à aplicação de FPAR.

Para exemplificar os resultados alcançados com o método proposto, a Figura 43 e a Figura 44 apresentam duas imagens distintas e seus respectivos resultados, além das imagens geradas para uso no processo de aplicação de textura.

Observa-se na Figura 43 que a máscara de textura que é multiplicada pelas imagens IRNA e IFPAR é de fundamental importância para a redução dos FP.

Figura 43 - Resultado do método proposto para a imagem: img (20).jpg.

Na Figura 44 pode-se observar que a aplicação do FPAR removeu a orelha do indivíduo, comprometendo a acurácia do processo, enquanto a segmentação por RNA manteve a orelha mas também um pedaço da roupa, que é posteriormente removido com a aplicação de textura.

Ao avaliar os resultados destas duas imagens, pode-se concluir que algumas imagens podem obter melhores resultados em uma abordagem, e outras obtêm melhor resultado em outra abordagem. Porém, de uma forma geral e considerando os valores de acurácia e falsos positivos, a aplicação de textura nas imagens resultantes do FPAR obtêm os melhores resultados considerando o objetivo deste trabalho.

Utilizando o resultado ótimo da aplicação de textura considerando o menor valor de FP alcançado, os gráficos exibidos na Figura 45 e na Figura 46 exibem o comparativo entre as abordagens de segmentação de pele utilizando RNA (IRNA), aplicação do FPAR (IFPAR), aplicação de textura nas IRNA (IRNA+TEX) e aplicação de textura nas IFPAR (IFPAR+TEX).

Figura 44 - Resultado do método proposto para a imagem: img (22).jpg.

Figura 45 - Comparativo das acurácias das aplicações.

Como abordado anteriormente ao longo do item 4.3, a acurácia não sofre alteração de forma significativa dentre as diversas aplicações, variando entre a maior acurácia alcançada (IFPAR) e a menor (IRNA+TEX) em 1,8% (Figura 45).

Figura 46 - Comparativo dos falsos positivos das aplicações.

Apesar da invariância nos valores de acurácia, os FP sofrem alteração mais significativa entre as diferentes aplicações. O maior valor de FP (IRNA) é de 11,42%. Este valor é reduzido para 6,06% (IFPAR+TEX), o que significa uma redução de 46,9%.

As conclusões do trabalho, suas principais contribuições e algumas sugestões para trabalhos futuros são apresentadas a seguir com base nos resultados apresentados neste capítulo.

5. CONCLUSÕES

A principal proposta deste trabalho foi o desenvolvimento de um método que permitisse a redução de falsos positivos em problemas de segmentação de pele humana aplicados a imagens faciais.

Com os resultados obtidos que são apresentados e discutidos no Capítulo 4 pode-se confirmar que este objetivo foi alcançado. Com o desenvolvimento de dois diferentes procedimentos aplicáveis às imagens faciais, pôde-se reduzir a quantidade de falsos positivos das imagens analisadas em até 46,9%.

A eficiência demonstrada pelo método proposto permite que, em casos onde há pixels com valores de cor condizentes aos valores de cor de pele e que não fazem parte do segmento de pele humana, se possa reduzir a classificação errônea destes pixels.

Os procedimentos abordados no método proposto neste trabalho encontram aplicabilidade direta em problemas que priorizem a menor quantidade de falsos positivos em relação a maior quantidade de verdadeiros positivos. Isto significa que os objetivos do desenvolvimento do trabalho foram atingidos, pois para aplicações onde é mais importante que não haja áreas classificadas como pele humana erroneamente, pode-se considerar a adoção dos procedimentos apresentados neste trabalho com segurança para a redução da frequência de falsos positivos. O método apresentado neste trabalho alcança resultados promissores atingindo até 46,9% de redução de falsos positivos sem que a acurácia aferida tenha redução elevada (apenas 1,8%).

Na abordagem onde há a aplicação do FPAR nas imagens segmentadas, além do percentual de falsos positivos ter redução significativa (42,2%), há também um aumento na acurácia da segmentação das imagens (0,7%). Esta abordagem pode ser utilizada em aplicações onde a acurácia do processo é mais relevante, ou aplicações onde não é necessária uma quantidade muito reduzida de falsos positivos.

Conclui-se com este trabalho um método que pode ser adaptado ao objetivo da aplicação, enfatizando a redução de falsos positivos na segmentação de pele humana em imagens faciais.

5.1. Contribuições

Dentre as contribuições deste trabalho, pode-se citar:

- Desenvolvimento de uma aplicação de reconhecimento facial utilizando apenas cores como descritores e que alcançou alto índice de acertos (em torno de 93%);
- Criação de um banco de imagens com amostras de pele e não pele que está disponível abertamente on-line e pode ser utilizado pela comunidade científica para desenvolvimento de métodos e aplicações ou mesmo para *benchmarking*;
- Desenvolvimento de um procedimento de fácil aplicabilidade que visa reduzir a quantidade de falsos positivos em problemas de segmentação de pele humana em imagens faciais (FPAR);
- Desenvolvimento de um procedimento de mais complexa aplicabilidade que permite reduzir a quantidade de falsos positivos da segmentação de pele humana em imagens faciais;

5.2. Sugestões para Trabalhos Futuros

As aplicações desenvolvidas neste trabalho não compreendem todo o espaço de possibilidades que o banco de imagens SFA e o próprio método permitem. Portanto, testes empíricos mais detalhados podem oferecer melhores resultados, como maior variação de *thresholds*, β e dimensões de amostras.

A aplicação deste método em outras bases de imagens ou até mesmo a aplicação deste adaptada a outros tipos de imagens, com objetivos diferenciados, são também alternativas de desenvolvimento futuro deste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

ANBARJAFARI, Gholamreza. **A new face recognition system based on colour statistics**. 2008. 64 f. Dissertação (Mestrado) - Curso de Master Of Science, Departamento de Electrical And Electronic Engineering, Eastern Mediterranean University, Gazimagusa, 2008.

BAKHSHALI, Mohamad Amin; SHAMSI, Mousa; GOLZARFAR, Amir. Facial Skin Segmentation with Varying Illumination Correction for Evaluation of Facial Surgery Systems. In: IEEE SYMPOSIUM ON INDUSTRIAL ELECTRONICS AND APPLICATIONS, 2012., Bandung. **Proceedings...** . Bandung: Ieee, 2012. p. 305 - 309.

BHATT, Rajen; DHALL, Abhinav. **Skin Segmentation Dataset**: UCI Machine Learning Repository. Disponível em: <<http://archive.ics.uci.edu/ml/datasets/Skin+Segmentation>>. Acesso em: 01 fev. 2013.

BIMBO, A. D. **Visual information retrieval**. Morgan Kaufmann Publishers Inc. San Francisco, California, 1999.

CASATI, João Paulo Brognoni; MORAES, Diego Rafael; RODRIGUES, Evandro Luis Linhari. SFA: A Human Skin Image Database based on FERET and AR Facial Images. In: IX WORKSHOP DE VISÃO COMPUTACIONAL, 9., 2013, Rio de Janeiro. **Anais...** . Rio de Janeiro, 2013. p. 1 - 5.

CASATI, João Paulo Brognoni; RODRIGUES, Evandro Luis Linhari. Comparação de Técnicas de Reconhecimento Facial Utilizando Histogramas de Cores Quantizadas pelo Misturograma. In: VIII WORKSHOP DE VISÃO COMPUTACIONAL, 8., 2012, Goiânia. **Anais...** . Goiânia, 2012. p. 1 - 5.

CELEBI, M. Emre; WEN, Quan; CHEN, Juan. Color Quantization Using C-Means Clustering Algorithms. In: IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 18., 2011, Bruxelas. **Proceedings...** . Bruxelas: Ieee, 2011. p. 1729 - 1732.

CHAVES-GONZÁLEZ, Jose M. et al. Detecting skin in face recognition systems: A colour spaces study. **Digital Signal Processing**, v. 20, n. 3, p.806-823, maio 2010.

CHOI, Jae Young; RO, Yong Man; PLATANIOTIS, Konstantinos N. Boosting Color Feature Selection for Color Face Recognition. **Ieee Transactions On Image Processing**, Piscataway, v. 20, n. 5, p.1425-1434, maio 2011.

GONZALEZ, Rafael C.; WOODS, Richard E. **Digital image processing**. Prentice Hall, 2008.

GONZALEZ, Rafael C.; WOODS, Richard E.; EDDINS, Steven L. **Digital Image Processing using Matlab**. 2. ed. Gatesmark Publishing, 2009. 827 p.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, v. SMC-3, n. 6, p.610-621, 1973.

HAYKIN, Simon. **Neural Networks: a comprehensive foundation**. 2. ed. Patparganj, Delhi: Prentice Hall, 1999. 842 p.

JAFRI, Rabia; ARABNIA, Hamid R.. A Survey of Face Recognition Techniques. **Journal Of Information Processing Systems**, Seoul, v. 5, n. 2, p.41-68, jun. 2009.

JIANG, Zhiwei; YAO, Min; JIANG, Wei. Skin Detection Using Color, Texture and Space Information. In: FOURTH INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY, 4., 2007, Haikou. **Proceedings...** . Haikou: Ieee, 2007. v. 3, p. 366 - 370.

JIN, Zhong et al. Face detection using template matching and skin-color information. **Neurocomputing**, S. L., v. 70, n. 4-6, p.794-800, jan. 2007.

JUSOH, Rizal Mat et al. Skin detection based on thresholding in RGB and hue component. In: IEEE SYMPOSIUM ON INDUSTRIAL ELECTRONICS & APPLICATIONS, 2010. Penang. **Proceedings...** . Penang: Ieee, 2010. p. 515 - 517.

KAKUMANU, P.; MAKROGIANNIS, S.; BOURBAKIS, N. A survey of skin-color modeling and detection methods. **Pattern Recognition**, v. 40, n. 3, p.1106-1122, mar. 2007.

KOSCHAN, Andreas; ABIDI, Mongi. **Digital color image processing**. Hoboken: John Wiley & Sons, 2008.

MARTINEZ, A. M.; BENAVENTE, R. **The AR Face Database**. Computer Vision Center Technical Report, n. 24, 1998.

MØLLER, Martin Fodslette. A scaled conjugate gradient algorithm for fast supervised learning. **Neural Networks**, v. 6, n. 4, p.525-533, 1993.

MORAES, Diego Rafael; CASATI, João Paulo Brognoni; RODRIGUES, Evandro Luis Linhari. Analysis of polynomial behavior of the C3 cervical concavity to bone age estimation using artificial neural networks. In: BIOSIGNALS AND BIOROBOTICS CONFERENCE, 4., 2013, Rio de Janeiro. **Proceedings...** . Rio de Janeiro: Ieee, 2013. p. 1 - 6.

NG, Pan; PUN, Chi-man. Skin Color Segmentation by Texture Feature Extraction and K-mean Clustering. In: THIRD INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE, COMMUNICATION SYSTEMS AND NETWORKS, 3., 2011, Bali. **Proceedings...** . Bali: Ieee, 2011. p. 213 - 218.

PELE, Ofir; WERMAN, Michael. The Quadratic-Chi Histogram Distance Family. In: 11TH EUROPEAN CONFERENCE ON COMPUTER VISION, 11., 2010, Crete. **Proceedings...** . Crete: Springer, 2010. p. 749 - 762.

PENATTI, Otávio Augusto Bizetto; TORRES, Ricardo da Silva. Color descriptors for Web image retrieval: a comparative study. In: BRAZILIAN SYMPOSIUM ON COMPUTER GRAPHICS AND IMAGE PROCESSING (SIBGRAPI'2008), 21., 2008, Campo Grande. **Proceedings...** . Campo Grande: Ieee, 2008. p. 163 - 170.

PERETTO, Pierre. **An Introduction to the Modeling of Neural Networks**. New York: Press Syndicate Of The University Of Cambridge, 1992. 492 p.

PHILLIPS, P. Jonathon et al. The FERET evaluation methodology for face recognition algorithms. **IEEE Transactions On Pattern Analysis And Machine Intelligence**, v. 22, n. 10, p.1090-1104, out. 2000.

QIAOQIAO, Zhu; YUANYUAN, Huang. A New Image Retrieval Method Based on Color Feature. In: INTERNATIONAL CONFERENCE ON INTELLIGENT SYSTEM DESIGN AND ENGINEERING APPLICATIONS (ISDA'2012), 12., 2012, Kochi. **Proceedings...** . Kochi: Ieee, 2012. p. 56 - 59.

SARKAR, Rajib; BAKSHI, Sambit; SA, Pankaj K. A Real-time Model for Multiple Human Face Tracking from Low-resolution Surveillance Videos. **Procedia Technology**, Rourkela, v. 1, n. 1, p.1004-1010, out. 2012.

SEVERINO JUNIOR, Osvaldo; GONZAGA, Adilson. Misturograma – Uma proposta de Quantização do Histograma através da Mistura de Cores. In: WORKSHOP DE VISÃO COMPUTACIONAL, 1., 2005, Piracicaba. **Anais...** . Piracicaba, 2005. p. 76 - 79.

_____. Auto-esboço e Auto-pele pelo Misturograma. In: WORKSHOP DE VISÃO COMPUTACIONAL (WVC'2006), 2., 2006, São Carlos. **Anais...** .São Carlos, 2006. p. 81 - 86.

_____. A new approach for color image segmentation based on color mixture. **Machine Vision And Applications**, S. L., v. 24, n. 3, p.607-618, abr. 2013.

SEVERINO JUNIOR, Osvaldo. **Mistura de cores: Uma nova abordagem para processamento de cores e sua aplicação na segmentação de imagens**. 2009. 121 f. Tese (Doutorado) - Universidade de São Paulo, São Carlos, 2009.

SHIH, Peichung; LIU, Chengjun. Extracting Efficient Color Features for Face Recognition Using Evolutionary Computation. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE AND MULTIMEDIA APPLICATIONS (ICCIMA'2005), 6., 2005, New Jersey. **Proceedings...** . New Jersey: IEEE, 2005. p. 1 - 6.

SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes Neurais Artificiais para Engenharia e Ciências Aplicadas: curso prático**. São Paulo: Artliber, 2010. 399 p.

SILVA, Leandro Sebastian Pereira da; GONZAGA, Adilson. Recuperação de Imagens de Faces Humanas Baseada em Conteúdo de retrato Falado. In: WORKSHOP DE VISÃO COMPUTACIONAL, 5., 2009, São Paulo. **Anais...** .São Paulo, 2009.

SIQUEIRA, Fernando Roberti de; SCHWARTZ, William Robson; PEDRINI, Helio. Adaptive Detection of Human Skin in Color Images. In: IX WORKSHOP DE VISÃO COMPUTACIONAL, 9., 2013, Rio de Janeiro. **Anais...** . Rio de Janeiro, 2013. p. 1 - 5.

VEZHNEVETS, Vladimir; SAZONOV, Vassili; ANDREEVA, Alla. A Survey on Pixel-Based Skin Color Detection Techniques. In: GRAPHICON 2003, 13., 2003, Moscou. **Proceedings...** . Moscou: Graphicon, 2003. p. 85 - 92.

VIEIRA, Raissa Tavares. **Análise de micropadrões em imagens digitais baseada em números fuzzy**. 2013. 151 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Universidade de São Paulo, São Carlos, 2013.

VIEIRA, Raissa Tavares et al. Image Micro-pattern Analysis Using Fuzzy Numbers. In: CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES (SIBGRAPI), 25., 2012, Ouro Preto. **Proceedings...** . Ouro Preto: Ieee, 2012a. p. 87 - 93.

_____. Local Fuzzy Pattern: A New Way for Micro-pattern Analysis. **Lecture Notes in Computer Science**. 1ed.: Springer Berlin Heidelberg, p. 602-611, 2012b.

_____. Uma nova proposta para análise de textura local por conjuntos fuzzy. In: VIII WORKSHOP DE VISÃO COMPUTACIONAL, 8., 2012, Goiânia. **Anais...** . Goiânia, 2012c. p. 1 - 5.

WANG, Baozhu; CHANG, Xiuying; LIU, Cuixiang. A Robust Method for Skin Detection and Segmentation of Human Face. In: SECOND INTERNATIONAL CONFERENCE ON INTELLIGENT NETWORKS AND INTELLIGENT SYSTEMS, 2009, 2., 2009, Tianjin. **Proceedings...** . Tianjin: Ieee, 2009. p. 290 - 293.

WANG, Xiaohua; ZHANG, Xi; YAO, Jingliang. Skin color detection under complex background. In: INTERNATIONAL CONFERENCE ON MECHATRONIC SCIENCE, ELECTRIC ENGINEERING AND COMPUTER., 2011, Jilin. **Proceedings...** . Jilin: Ieee, 2011. p. 1985 - 1988.

APÊNDICE A – EXEMPLOS DE IMAGENS DO SFA

O banco de imagens SFA é descrito no item 3.2 e está disponível em sua totalidade para download no endereço: <http://www.sel.eesc.usp.br/sfa>.

Alguns exemplos de suas imagens e amostras (apenas as de dimensão 35x35, pois englobam as outras dimensões) são apresentadas abaixo:

Arquivo	Imagem Original	Ground Truth	Amostras: Pele	Amostras: Não Pele
img (1).jpg
img (4).jpg

img (11).jpg
img (133).jpg
img (142).jpg
img (198).jpg

img (243).jpg
img (588).jpg
img (731).jpg
img (946).jpg

<p>img (1040).jpg</p>
<p>img (1044).jpg</p>
<p>img (1093).jpg</p>

APÊNDICE B – CÓDIGO-FONTE: TRANSFORMAÇÃO DO RGB PARA RGB NORMALIZADO

Segue abaixo o código-fonte da função que transforma uma imagem do espaço de cores RGB para o espaço de cores RGB Normalizado.

Implementação para MATLAB (arquivo “*rgb2normrgb.m*”):

```
function [ norm ] = rgb2normrgb( f )
%RGB2NORMRGB - This script receives a RGB image and
returns a NormRGB.
%image.
% Implemented by João Paulo Brognoni Casati (2013).

    im = double(f);

    R = im(:,:,1);
    G = im(:,:,2);
    B = im(:,:,3);

    NormalizedRed = R(:,:,:)./(R(:,:,:)+G(:,:,:)+B(:,:,:));
    NormalizedGreen = G(:,:,:)./(R(:,:,:)+G(:,:,:)+B(:,:,:));
    NormalizedBlue = B(:,:,:)./(R(:,:,:)+G(:,:,:)+B(:,:,:));

    norm(:,:,1) = NormalizedRed(:,:,:);
    norm(:,:,2) = NormalizedGreen(:,:,:);
    norm(:,:,3) = NormalizedBlue(:,:,:);
end
```


APÊNDICE C – CÓDIGO-FONTE: PROCEDIMENTOS DO MÉTODO PROPOSTO

Segue abaixo o código-fonte do script que executa os procedimentos do método nas 1118 imagens do banco SFA.

Implementação para MATLAB (arquivo “*main.m*”):

```
% Procedimentos para segmentação de pele
% Implemented by João Paulo B. Casati
%

% function confusion2() => returns accuracy and
% confusion matrix.

% function img2vec() => transforms image in vector

% function proc_img_rna() => returns ann output

% function proc_img_chi() => returns chi-square distance

close all;
clear all;
clc;

% Init Variables

% weights bias:
% best training (samples)
load('net_melhor_rgb');

% samples size
vec_am = [5 11 15 25 35];

% skin mean histograms
load('skin_mean_h5');
load('skin_mean_h11');
load('skin_mean_h15');
load('skin_mean_h25');
load('skin_mean_h35');
skin(5).h = skin_mean_h5;
skin(11).h = skin_mean_h11;
skin(15).h = skin_mean_h15;
skin(25).h = skin_mean_h25;
skin(35).h = skin_mean_h35;
```

```

masc = 3;

for i=1:1118

% Image segmentation
%
f = imread(['C:/imagens/ORI/img (' int2str(i) ').jpg']);
f = imresize(f,0.25);
[x y z] = size(f);

% ground truths already binarized
gt = imread(['C:/imagens/GTBW/img (' int2str(i) ').jpg']);
gv = img2vec(im2bw(gt));

% ANN Segmentation
%
[rnai rnav] = proc_img_rna(f,net,0.1);

[a conf] = confusion2(gv,rnav);
res_rna(i).conf = conf';
res_rna(i).a = (1-a)*100;

% FPAR Segmentation
%
nk = 0;
for k=0.05:0.05:0.95
    num = sum(sum(rnai));
    perc = num*k;
    fpri = bwareaopen(rnai,round(perc));

    fprv = img2vec(fpri);
    [a conf] = confusion2(gv,fprv);
    nk = nk+1;
    res_fpr(nk,i).conf = conf';
    res_fpr(i).a = (1-a)*100;
end

% TEX Segmentation
% k = 0.15 (result of FPAR)

nbeta = 0;
for beta=0.05:0.05:2
    nbeta = nbeta+1;

    lfp = lfp(rgb2gray(f),masc,beta);
    lfpi = zeros(x,y);
    lfpi(2:x-1,2:y-1) = lfp;
    clear lfp;
    for nam=1:length(vec_am)

        amostra = vec_am(nam);

        fator = ceil(amostra/2);

        disti = proc_img_chi(lfpi,amostra,skin(amostra).h);
    end
end

```

```

semborda = disti(fator+1:x-fator,fator+1:y-fator);

img_tex = ones(x,y);

img_tex(fator+1:x-fator,fator+1:y-fator) = mat2gray(semborda);

nth = 0;
for th=0.05:0.05:0.95
    nth = nth+1;

    masc_tex = imcomplement(im2bw(img_tex,th));

    %RNA + TEX
    rna_tex_img = masc_tex .* rnai;
    rna_tex_vec = img2vec(rna_tex_img);

    %RNA + FPR + TEX
    fpr_tex_img = masc_tex .* fpri;
    fpr_tex_vec = img2vec(fpr_tex_img);

    %RESULTADOS
    res_tex(i,nbeta,nam,nth).beta = beta;
    res_tex(i,nbeta,nam,nth).amostra = amostra;
    res_tex(i,nbeta,nam,nth).th = th;

    [a conf] = confusion2(gv,rna_tex_vec);
    res_tex(i,nbeta,nam,nth).confrnatex = conf';
    res_tex(i,nbeta,nam,nth).arnatex = (1-a)*100;

    [a conf] = confusion2(gv,fpr_tex_vec);
    res_tex(i,nbeta,nam,nth).conffprt看 = conf';
    res_tex(i,nbeta,nam,nth).afprt看 = (1-a)*100;

end
end
end
end

```