

CENTRO UNIVERSITÁRIO FEI
PEDRO AUGUSTO SANTOS ORONA SILVA

**INTERPRETAÇÃO E RECONHECIMENTO DE PADRÕES PARA AVALIAÇÃO DE
DOR EM IMAGENS FACIAIS DE RECÉM-NASCIDOS**

São Bernardo do Campo

2020

PEDRO AUGUSTO SANTOS ORONA SILVA

**INTERPRETAÇÃO E RECONHECIMENTO DE PADRÕES PARA AVALIAÇÃO DE
DOR EM IMAGENS FACIAIS DE RECÉM-NASCIDOS**

Dissertação de Mestrado, apresentada ao Centro
Universitário FEI para obtenção do título de Mes-
tre em Engenharia Elétrica. Orientada pelo Prof. Dr.
Carlos Eduardo Thomaz.

São Bernardo do Campo

2020

Silva, Pedro Augusto Santos Orona.

Interpretação e reconhecimento de padrões para avaliação de dor em imagens faciais de recém-nascidos / Pedro Augusto Santos Orona Silva. São Bernardo do Campo, 2020.

85 p. : il.

Dissertação - Centro Universitário FEI.

Orientador: Prof. Dr. Carlos Eduardo Thomaz.

1. Processamento de Imagem. 2. Imagem Facial. 3. Bebês a termo. 4. Extração de características. 5. Reconhecimento de padrões. I. Thomaz, Carlos Eduardo, orient. II. Título.

Aluno: Pedro Augusto Santos Orona Silva

Matrícula: 118119-7

Título do Trabalho: Interpretação e reconhecimento de padrões para avaliação de dor em imagens faciais de recém-nascidos.

Área de Concentração: Processamento de Sinais e Imagens

Orientador: Prof. Dr. Carlos Eduardo Thomaz

Data da realização da defesa: 28/02/2020

ORIGINAL ASSINADA

Avaliação da Banca Examinadora:

São Bernardo do Campo, / / .

MEMBROS DA BANCA EXAMINADORA

Prof. Dr. Carlos Eduardo Thomaz

Ass.: _____

Prof. Dr. Paulo Sergio Silva Rodrigues

Ass.: _____

Profª Drª Maria Fernanda Branco de Almeida

Ass.: _____

A Banca Julgadora acima-assinada atribuiu ao aluno o seguinte resultado:

APROVADO

REPROVADO

VERSÃO FINAL DA DISSERTAÇÃO

**APROVO A VERSÃO FINAL DA DISSERTAÇÃO EM QUE
FORAM INCLUÍDAS AS RECOMENDAÇÕES DA BANCA
EXAMINADORA**

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Carlos Eduardo Thomaz

À minha família e à todos que acreditaram nesse sonho.

AGRADECIMENTOS

À minha família como um todo, principalmente aos meus pais, João Ayres e Kátia Celene, por todo o suporte e encorajamento oferecido por eles ao longo do desenvolvimento deste trabalho, sempre me mostrando que a vida é justa e os frutos são grandiosos caso exista esforço e dedicação.

Agradeço de todo coração ao meu orientador, Dr. Carlos Eduardo Thomaz, por ter me recebido de tão longe e acreditado no meu potencial. Sou imensamente grato por todos os ensinamentos que me foram repassados, pela paciência que o senhor teve comigo durante todas as nossas discussões, e pela oportunidade dada.

Aos meus amigos de mestrado, em especial à M.^a Estela Ribeiro e ao M.e Davi Araujo, que foram como mestres para mim, me ajudando de todas as maneiras possíveis, mostrando pontos de falha na minha redação e auxiliando em questões técnicas. Muito obrigado, sem vocês nada disso seria possível.

Aos meus amigos e colegas, que sempre me apoiaram e acreditaram neste sonho, ouvindo minhas angústias e felicidades ao longo deste trabalho. Agradeço em especial aos amigos de Campo Grande, Rafael e Júlio Marcos, e aos meus amigos de república, Gabriel, Guilherme, Jéssica e Vinícius. Foi muito bom saber que vocês estavam do meu lado.

Ao BTG Pactual pela ajuda e suporte, em especial ao Felipe Braga e ao Estevam Carvalho, que me acolheram em um ambiente totalmente novo para mim e me apoiaram na conclusão deste trabalho.

Finalizo agradecendo à CAPES e ao Centro Universitário FEI pela bolsa de estudos do mestrado.

“Eu acredito que às vezes são as pessoas de quem ninguém espera nada que fazem as coisas que ninguém consegue imaginar.”

Alan Turing

RESUMO

Analisar a dor em recém-nascidos se tornou um estudo de relevância nas últimas décadas dada a impossibilidade de se conseguir indicar objetivamente qual a localização e intensidade deste fenômeno doloroso. Ao longo dos últimos anos, métodos de detecção e avaliação de dor foram capazes de classificar dor através de imagens faciais de bebês recém-nascidos, utilizando modelos estatísticos, aprendizado de máquina e aprendizado profundo. Neste contexto, há interesse dos profissionais da saúde em terem disponível ferramentas computacionais capazes de explicitar não somente o nível potencial de dor experienciado pelo recém-nascido, mas também as regiões faciais de maior relevância para o fenômeno dor. O objetivo desta dissertação é desenvolver um arcabouço computacional de interpretação e reconhecimento de padrões em imagens de faces para avaliação automática de dor em bebês a termo. Mais especificamente, este trabalho concentra-se na investigação, implementação e integração de técnicas de detecção, segmentação, normalização espacial e classificação de imagens de faces baseadas em informações extraídas por mineração estatística de dados. Por fim, o arcabouço desenvolvido aqui, avaliado com uma taxa de acerto (limite superior) de aproximadamente 96% para a base COPE e 77% para a base UNIFESP, mostra que é possível não somente classificar estatisticamente dor e não-dor através de imagens de faces, mas também evidenciar regiões faciais discriminantes para o fenômeno dor, auxiliando na construção de escalas pediátricas de dor mais gerais e assertivas.

Palavras-chave: Processamento de Imagem. Imagem Facial. Bebês a termo. Extração de características. Reconhecimento de padrões. Dor.

ABSTRACT

Pain analysis in newborns has become a relevant study subject over the last few decades, given the inability to objectively identify the source and intensity of the pain in newborn babies. Over the last few years, several methods for pain detection and evaluation were able to classify pain levels using facial expressions from newborn babies, through statistical models, machine learning and deep learning. Considering this context, health professionals are increasingly more interested in having computerized tools at their disposal. These tools would not only be able to accurately rank the newborn's potential pain level, but also identify the facial regions of greatest relevance for a particular pain phenomenon. This dissertation's main objective is to develop a computer framework capable of recognizing and interpreting patterns in facial expressions for an automated evaluation of pain levels on term babies. Specifically, this dissertation focuses on the investigation, implementation and integration of a series of techniques, including image detection and segmentation, spacial normalization and, ultimately, the classification of facial expressions based on information obtained through statistical data mining. Finally, the framework developed here, evaluated with an accuracy (upper limit) of approximately 96% for the COPE base and 77% for the UNIFESP base, reveal that it is possible to not only rank pain levels statistically through images of facial expressions, but also to identify key facial regions for certain pain phenomena, therefore assisting in creating more general and accurate pediatric pain scales.

Keywords: Image processing. Face images. Term babies. Feature extraction. Pattern recognition. Pain.

LISTA DE FIGURAS

Figura 1 – Dois conjuntos de exemplos das cinco expressões presentes na base de dados COPE	23
Figura 2 – Exemplos da base de faces COPE	33
Figura 3 – Exemplos da base de dados NPAD-1	35
Figura 4 – Posicionamento das três câmeras envolta do berço	36
Figura 5 – Exemplo de imagem do relatório (esquerda) e da imagem avaliada no relatório (direita) utilizada em Heiderich, Leslie e Guinsburg (2015)	37
Figura 6 – Exemplos da base de dados UNIFESP	38
Figura 7 – Etapas para normalização global das imagens: (a) marcação automática dos 68 pontos fiduciais na imagem original preprocessada, (b) especificação das 5 coordenadas de destino (olhos, nariz e cantos da boca) e (c) a imagem normalizada globalmente	42
Figura 8 – Etapas para normalização não-rígida das imagens: (a) Malha triangular de saída, (b) especificação da malha triangular de destino, (c) a imagem correspondente à malha de saída e (d) a imagem resultante normalizada não rigidamente.	43
Figura 9 – Processo de normalização utilizando todos-contra-todos	44
Figura 10 – Atlas das imagens da COPE: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente	53
Figura 11 – Atlas das imagens da UNIFESP: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente	54
Figura 12 – Desvio padrão dos atlas das imagens da COPE: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente	56
Figura 13 – Desvio padrão dos atlas das imagens da UNIFESP: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente	57
Figura 14 – Projeção dos dados no eixo discriminante entre o grupo com dor (esquerda) e sem dor (direita) para a base COPE	60
Figura 15 – Navegação no hiperplano das sete primeiras componentes principais, ordenadas da maior para menor variância, para a base COPE	61
Figura 16 – Navegação no hiperplano discriminante para a base COPE	62

Figura 17 – Projeção dos dados no eixo discriminante entre o grupo com dor e sem dor removendo a primeira componente principal para a base COPE	62
Figura 18 – Navegação no hiperplano discriminante removendo a primeira componente principal original do PCA para a base COPE	63
Figura 19 – Projeção dos dados no eixo discriminante entre o grupo com dor (esquerda) e sem dor (direita) para a base UNIFESP	63
Figura 20 – Navegação no hiperplano das sete primeiras componentes principais, ordenadas da maior para menor variância, para a base UNIFESP	64
Figura 21 – Navegação no hiperplano discriminante para a base UNIFESP	65
Figura 22 – Projeção dos dados no eixo discriminante entre o grupo com dor e sem dor removendo a primeira componente principal para a base UNIFESP	65
Figura 23 – Navegação no hiperplano discriminante removendo a primeira componente principal original do PCA para a base COPE	66
Figura 24 – K-fold com $k = 5$ utilizando o primeiro hiperplano discriminante construído para a base COPE	67
Figura 25 – K-fold com $k = 5$ utilizando o segundo hiperplano discriminante construído para a base COPE	68
Figura 26 – K-fold com $k = 5$ utilizando o primeiro hiperplano discriminante construído para a base UNIFESP.	69
Figura 27 – K-fold com $k = 5$ utilizando o segundo hiperplano discriminante construído para a base UNIFESP	70
Figura 28 – Seleção de pares para construção dos mapas	72
Figura 29 – Mapas T das imagens geradas na navegação do hiperplano do MLDA para a base COPE: (a) mapa 1, (b) mapa 2, (c) mapa 3 e (d) mapa 4	73
Figura 30 – Mapas T das imagens geradas na navegação do hiperplano do MLDA para a base UNIFESP (a) mapa 1, (b) mapa 2, (c) mapa 3 e (d) mapa 4	74
Figura 31 – Regiões de destaque para a base COPE	75
Figura 32 – Regiões de destaque para a base UNIFESP	76

LISTA DE TABELAS

Tabela 1 – Escalas pediátricas de dor	19
Tabela 2 – Métodos automáticos para análise de expressões de dor em recém-nascidos .	21
Tabela 3 – Extração de características para análise de expressões de dor em recém-nascidos	22
Tabela 4 – Semelhança entre as imagens da base COPE	58
Tabela 5 – Semelhança entre as imagens da base da UNIFESP	58
Tabela 6 – Resultados das classificações para a base COPE	71
Tabela 7 – Resultados das classificações para a base UNIFESP	71

LISTA DE ABREVIATURAS

AUC	<i>Area Under The Curve.</i>
CNN	Redes Neurais Convolucionais.
COPE	Classificação de Expressões de Dor.
EC	Entropia Conjunta.
ELBP	Padrão Binário Alongado.
ELTP	Padrão Ternário Alongado.
IM	Informação Mútua.
IMN	Informação Mútua Normalizada.
LBP	Padrões Binários Locais.
LBP-TOP	Padrões Binários Locais em Três Planos Ortogonais.
LDA	Análise Discriminante Linear.
LDP	Padrão Direcional Local.
LDTP	Padrão de Textura Direcional Local.
LGC	Código de Gradiente Local.
LTP	Padrão Ternário Local.
NFCS	Sistema de Código Facial Neonatal.
NPAD	Neonatal Pain Assessment Dataset.
PCA	Análise das Componentes Principais.
RVM	Máquina de Vetores de Relevância.
SFFS	Sequential Forward Floating Selection.
SVM	Máquina de Vetores de Suporte.
WLBP	Padrão Binário Local Ponderado.

SUMÁRIO

1	INTRODUÇÃO	14
1.1	OBJETIVO	15
2	CONCEITOS FUNDAMENTAIS E TRABALHOS RELACIONADOS	17
2.1	ESCALAS PEDIÁTRICAS DE DOR	17
2.2	CLASSIFICAÇÃO DE DOR POR EXPRESSÕES FACIAIS	20
2.2.1	Métodos Baseados em Redução de Características	22
2.2.2	Métodos Baseados em Variação de Padrões Binários Locais	26
2.2.3	Aprendizado Profundo	29
2.3	RECONHECIMENTO DE DOR EM IMAGENS FACIAIS	29
2.3.1	Detecção de dor	29
2.3.2	Estimativa de intensidade de dor	30
2.4	BASES DE DADOS	30
2.4.1	COPE	31
2.4.2	NPAD	33
2.4.3	UNIFESP	35
2.5	CONSIDERAÇÕES COMPLEMENTARES	38
3	METODOLOGIA	39
3.1	NORMALIZAÇÃO ESPACIAL DAS IMAGENS	39
3.1.1	Pré-processamento	39
3.1.2	Predição dos pontos fiduciais da face	40
3.1.3	Alinhamento global	40
3.1.4	Alinhamento não-rígido	42
3.2	NORMALIZAÇÃO TODOS CONTRA TODOS	43
3.3	EXTRAÇÃO MULTIVARIADA DE INFORMAÇÃO	45
3.3.1	Análise das Componentes Principais (PCA)	45
3.3.2	Análise Linear Discriminante (LDA)	46
3.4	MÉTRICAS DE SIMILIARIDADE	48
3.4.1	Minimização do Erro Quadrático Médio	48
3.4.2	Coefficiente de Correlação de Pearson	49
3.4.3	Informação Mútua Normalizada	49
3.4.4	Norma L^2	50

3.5	MAPAS T	51
4	EXPERIMENTOS E RESULTADOS	52
4.1	ATLAS	52
4.2	IMAGENS DE ERRO (QUALITATIVA)	55
4.3	MÉTRICAS DE SIMILARIDADE (QUANTITATIVA)	57
4.4	EXTRAÇÃO MULTIVARIADA DE INFORMAÇÃO	58
4.4.1	SELEÇÃO DAS IMAGENS	59
4.4.2	CONSTRUÇÃO DO HIPERPLANO	59
4.5	VALIDAÇÃO	66
4.6	MAPAS T DE DOR	72
5	CONCLUSÃO	77
	REFERÊNCIAS	79

1 INTRODUÇÃO

A dor é uma característica intrínseca a qualquer ser vivo que possua sistema nervoso central (ANAND; CRAIG, 1996), servindo para alertar que algo está errado, uma presença de dano corporal ou ameaça de morte (LUDA et al., 2007). Clinicamente, a dor é definida como uma experiência subjetiva e o método mais confiável para avaliá-la é através do auto-relato, ou seja, por meio de expressões e indicações pontuais individuais (LUDA et al., 2007).

Para a maioria dos adultos é fácil descrever verbalmente a localização, duração e intensidade de suas experiências dolorosas, porém para o caso de recém-nascidos, é preciso uma atenção maior, dada a impossibilidade contextual de indicação e classificação de dor. Esta necessidade se torna ainda mais importante dado o fato de que pesquisas recentes indicam que dores não tratadas em recém-nascidos podem resultar em mudanças no sistema nervoso central que desaceleram o desenvolvimento dos indivíduos (LUDA et al., 2007). Logo, para o caso dos recém-nascidos, há exigência de serem desenvolvidos métodos de análise que busquem classificar a dor utilizando informações fisiológicas e comportamentais (GUINSBURG; CUENCA, 2010).

Com este objetivo, diversas escalas de dor foram construídas pela comunidade de neonatologia, tendo como parâmetros posições de pontos específicos da face, movimentação do corpo e etc (GUINSBURG, 1999; GUINSBURG; CUENCA, 2010). Porém, existem dificuldades em utilizá-las principalmente por não haver uma escala global, que sirva como padrão para qualquer avaliação. A partir desta dificuldade, pesquisas surgiram com o objetivo de investigar e criar métodos de avaliação não-invasivos e mais específicos do fenômeno doloroso, que podem possibilitar a avaliação constante ou em períodos específicos do recém-nascido, auxiliando não somente na identificação precoce do fenômeno, como também em uma melhor compreensão da experiência dolorosa enfrentada pelo recém-nascido (HEIDERICH; LESLIE; GUINSBURG, 2015; BRAHNAM et al., 2006). Seguindo esta premissa, por meio da inteligência artificial e da visão computacional, trabalhos utilizando análise automática de dor através de imagens faciais de recém-nascidos (LU et al., 2016; ZAMZMI et al., 2018; ZHI et al., 2018) começaram a surgir, principalmente pela carência de monitoramentos de dor contínuos e consistentes em clínicas e centros de saúde, mostrando que talvez seja possível abordar a dificuldade de se classificar dor em recém-nascidos como um problema computacionalmente solucionável.

Tomando-se como referência o trabalho desenvolvido por Heiderich, Leslie e Guinsburg (2015), o qual trouxe como objetivo principal desenvolver e validar um software de computa-

dor capaz de monitorar movimentos faciais de dor em recém nascidos, esta dissertação propôs como primeiro passo, reunir novas referências, buscando trabalhos relacionados que procuram também reconhecer dor em recém-nascidos utilizando métodos computacionais. Através desta revisão da literatura, foi possível perceber que o tópico de detecção automática de dor em recém-nascidos é um assunto emergente, que gera desafios para a área de computação e também traz à tona novos questionamentos à medida que novos achados surgem.

A partir de trabalhos que procuram reconhecer dor em recém-nascidos, tendo como maior referência e ponto de início os trabalhos desenvolvidos por Teruel (s.d.) e Teruel et al. (2019), esta dissertação tem como objetivo propor e implementar, através do pré-processamento e análise multivariada, um método não invasivo capaz de identificar, classificar e avaliar dor utilizando imagens faciais de bebês a termo. Foi construído aqui um arcabouço computacional baseado em detecção e segmentação de objetos, normalização espacial e classificação de imagens de faces para caracterização e quantificação multivariada (THOMAZ; KITANI; GILLIES, 2006; SATO et al., 2008; THOMAZ; GIRALDI, 2010; THOMAZ et al., 2013; XAVIER et al., 2015, 2016a) de diferenças estatisticamente significantes, tendo a etapa de normalização como um dos processos obrigatórios para a classificação de dor, a qual não foi implementada em outros trabalhos que buscam o mesmo objetivo de classificar dor utilizando imagens faciais de bebês recém-nascidos.

Estes métodos foram aplicados em duas bases de imagens distintas, uma delas disponibilizada pela UNIFESP (HEIDERICH; LESLIE; GUINSBURG, 2015), e outra disponível publicamente via solicitação, denominada COPE (BRAHNAM et al., 2006). Ambas as bases foram construídas em ambiente controlado e são compostas por imagens faciais de bebês a termo que vivenciaram algum tipo ou nível de dor em decorrência de procedimentos clínicos tradicionais e necessários para diagnóstico e/ou tratamento de morbidades próprias do período neonatal (HEIDERICH; LESLIE; GUINSBURG, 2015).

Os resultados experimentais desta dissertação mostram que é possível não somente classificar estatisticamente dor e não-dor através de imagens de face, mas também evidenciar regiões faciais discriminantes para o fenômeno dor, auxiliando na construção de escalas pediátricas de dor mais gerais e assertivas.

1.1 OBJETIVO

O principal objetivo deste trabalho é desenvolver um arcabouço computacional capaz de avaliar dor automaticamente utilizando imagens faciais de bebês a termo, mas especificamente, é proposto um processo bem definido de normalização, extração de características de dor e classificação binária (dor ou não-dor). Sendo que a etapa de normalização se dá como uma etapa importante, dado que a maioria das bases de imagens disponíveis hoje apresentam um conjunto de dados não-normalizados (iluminação e posicionamento do bebê não controlados, entre outros).

Para este trabalho, o maior foco se encontra em explorar, quantificar e determinar medidas comportamentais baseadas em análise estatística multivariada de dor utilizando imagens de expressões faciais. Portanto, mais importante ainda do que oferecer uma aplicação que classifique imagens faciais de recém-nascidos como dor ou não-dor, este trabalho oferece aos profissionais de saúde modelos generalistas que podem auxiliá-los na tomada de decisões terapêuticas, indicando mais um sinal, ou padrão, de que determinado bebê está ou não experienciando algum tipo de dor. Semelhante ao papel das escalas pediátricas de dor no processo de avaliação de dor, os modelos propostos aqui pretendem oferecer um aporte a mais aos profissionais da saúde.

Espera-se também, por meio deste trabalho, estimular o uso de métodos de avaliação automática de dor em recém-nascidos na prática clínica, uma vez que se dispõe de uma metodologia não-invasiva de processamento e classificação de imagens para acompanhamento em tempo real do recém-nascido.

2 CONCEITOS FUNDAMENTAIS E TRABALHOS RELACIONADOS

Este capítulo descreve sucintamente conceitos fundamentais e os principais trabalhos relacionados aos assuntos estudados nesta dissertação de mestrado.

2.1 ESCALAS PEDIÁTRICAS DE DOR

A habilidade limitada em um recém-nascido de moderar sua experiência de dor enfatiza a extrema importância na avaliação realizada por profissionais da saúde, determinando se há sofrimento por parte do bebê (ANAND; STEVENS; MCGRATH, 2007). A avaliação precisa de dor é vital pois auxilia profissionais a entenderem a severidade da situação de seus pacientes e desenvolverem tratamentos adequados.

Segundo Anand e Craig (1996), "a percepção de dor é uma qualidade inerente à vida e manifesta-se já no início do desenvolvimento do indivíduo, servindo como um sistema de sinalização associado a lesões teciduais". Portanto, é indispensável que profissionais de saúde tenham condições de avaliar dor em recém-nascidos.

Para que seja possível atuar de forma terapêutica diante de situações possivelmente dolorosas, não basta apenas saber se o recém-nascido possui condições de exprimir a dor (GUINSBURG; CUENCA, 2010). É preciso, também, dispor de ferramentas que "decodifiquem" essa linguagem da dor. Dada esta necessidade, foram desenvolvidas escalas unidimensionais, que avaliam a resposta comportamental à dor, e instrumentos multidimensionais, que incluem uma combinação de parâmetros objetivos e subjetivos, empregados para avaliar componentes sensoriais, afetivos e avaliativos, os quais estão refletidos na linguagem utilizada pelo recém-nascido para descrever a experiência dolorosa (SCOPEL; ALENCAR; CRUZ, 2007; GUINSBURG; CUENCA, 2010).

Em Melo et al. (2014), foi realizada uma análise da literatura com o intuito de identificar instrumentos de avaliação da dor em recém-nascidos, sendo que a maioria dos artigos selecionados classifica as escalas segundo o tipo de Dor Aguda e Prolongada/Crônica. Tal classificação considera que a dor aguda frequentemente é causada por estímulos nociceptivos resultantes de lesões teciduais resultantes de procedimentos ou acidentes, porém costuma desaparecer à medida que ocorre cicatrização da ferida. Em casos de dor prolongada ou crônica, frequentemente, ocorre processo inflamatório deflagrado ou em continuidade a um fenômeno doloroso agudo,

sendo que muitos consideram crônica a dor com duração superior a 3 meses, ou seja, acima do período neonatal de 28 dias (OKADA et al., 2001).

Mesmo por meio da utilização das escalas de dor, a avaliação em recém-nascidos representa um grande desafio. As principais limitações para alcançar alívio da dor na prática clínica incluem a falta de um biomarcador confiável e a ausência de uma escala padrão ouro que seja capaz de medir a intensidade da dor, a necessidade de tratamento e a eficácia da intervenção (ARIAS; GUINSBURG, 2012). Vale ressaltar que, para alguns especialistas, o uso de escalas multidimensionais em recém-nascidos é mais adequado, por avaliarem repostas comportamentais associadas às respostas fisiológicas à dor, tornando a abordagem mais completa possível, tendo em vista que o relato de dor não pode ser manifestado nesse público (MELO et al., 2014). Já em Guinsburg e Cuenca (2010), relata-se que as escalas unidimensionais parecem ser ferramentas mais apropriadas para identificar os indivíduos com dor, quando comparadas às escalas multidimensionais.

Nesse contexto, indica-se, para avaliação da dor neonatal, o emprego de múltiplas escalas pelos diferentes profissionais de saúde, porém recomenda-se que dentro desse conjunto esteja presente pelo menos uma das escalas unidimensionais comportamentais (GUINSBURG; CUENCA, 2010). A Tabela 1 reúne as escalas pediátricas mais comuns utilizadas para diferentes tipos de dor.

Tabela 1 – Escalas pediátricas de dor

Escala de Dor	Classificação	Tipo de dor	Faixa etária	Medidas comportamentais	Medidas fisiológicas
<i>Neonatal Facial Coding System (NFCS)</i> (GRUNAU; CRAIG, 1987)	Unidimensional	Dor aguda e prolongada	Prematuro \geq 25 semanas de gestação a bebês a termo	Protuberância da sobrancelha, aperto nos olhos, sulco nasolabial, lábios abertos, boca horizontal, boca vertical, lábios franzidos, língua esticada, queixo tremendo, protrusão da língua	Nenhuma
COMFORT (AMBUEL et al., 1992)	Multidimensional	Dor aguda e prolongada	Recém nascidos pré-termo e de termo; até aos 3 anos	Alerta, agitação, reação respiratória, choro, movimentos gerais, tônus muscular e expressão facial	Pressão arterial e frequência cardíaca
<i>Neonatal Infant Pain Scale (NIPS)</i> (LAWRENCE et al., 1993)	Multidimensional	Dor aguda	28-38 semanas de gestação	Expressão facial, choro, braços/pernas, movimento e estado de excitação	Padrões de respiração
<i>Infant Body Coding System (IBCS)</i> (CRAIG et al., 1993)	Multidimensional	Dor aguda	Recém nascidos de termo e pré-termo	Expressão facial, movimentos corporais, características do choro	Frequência cardíaca e sinais vitais
<i>Crying, Requires O₂, Increased VS, Expression, and Sleepless (CRIES)</i> (KRECHEL; BILDNER, 1995)	Multidimensional	Dor prolongada	32-60 semanas de gestação	Expressão facial, choro, e estado adormecido	Oxigênio para saturação acima de 90% e aumento de sinais vitais
<i>Premature Infant Pain Profile (PIPP)</i> (STEVENS et al., 1996)	Multidimensional	Dor aguda e prolongada	Bebês com 37 semanas de gestação ou menos	Testa franzida, olhos espremidos e aprofundamento do sulco nasolabial	Frequência cardíaca e saturação de oxigênio
<i>Faces Leg Activity Cry Consolability Pain Scale (FLACC)</i> (MERKEL et al., 1997)	Multidimensional	Dor aguda e prolongada	Bebês com 1 ano ou mais que não conseguem relatar dor	Expressão facial, pernas, atividade, choro, e consolo	Nenhuma
<i>Scale for use in Newborns (SUN)</i> (BLAUER; GERSTMANN, 1998)	Multidimensional	Dor aguda	Recém nascidos de termo e pré-termo	Estado de alerta, tônus, movimento, expressão facial	Frequência cardíaca, respiração e pressão arterial média
<i>Visual Analog Scale (VAS)</i> (BUCHHOLZ et al., 1998)	Unidimensional	Dor aguda e prolongada	Recém nascidos de termo e pré-termo, até aos 4 anos	Expressão Facial	Nenhuma
<i>Échelle Douleur Inconfort Nouveau-Né (EDIN)</i> (DEBILLON et al., 2001)	Unidimensional	Dor aguda	25-36 semanas de gestação	Expressão facial, movimento corporal, qualidade do sono, contato com enfermagem, e consolabilidade	Nenhuma
<i>Escala de dor ABC</i> (BELLIENI et al., 2005)	Unidimensional	Dor aguda	Recém nascidos de termo e pré-termo	Tonalidade do primeiro choro, ritmo dos acessos de choro, constância da intensidade do choro	Nenhuma
<i>Pain Assessment Tool (PAT)</i> (SPENCE et al., 2005)	Multidimensional	Dor prolongada	Recém nascidos de termo e pré-termo	Postura/tônus, padrão de sono, expressão, coloração, choro	Frequência cardíaca, saturação do oxigênio, pressão arterial, respiração e percepção da dor pelo(a) enfermeiro(a)
<i>Behavioral Indicators of Infant Pain (BIIP)</i> (HOLSTI; GRUNAU, 2007)	Unidimensional	Dor aguda	23-32 semanas de gestação	Fronte saliente, olhos espremidos, sulco nasolabial aprofundado, estiramento horizontal da boca, língua tensa, mão espalmada, mão fechada e estado de sono/vigília	Nenhuma
<i>Neonatal Pain, Agitation, and Sedation Scale (N-PASS)</i> (HUMMEL et al., 2008)	Multidimensional	Dor prolongada	23-40 semanas de gestação	Expressão facial, movimentos de comportamento, choro/irritabilidade, e tônus nas extremidades	Frequência Cardíaca, frequência respiratória, pressão sanguínea, e saturação de oxigênio

Fonte: Autor “adaptado de” Zamzmi et al., 2017, p. 2; Da Silva e Da Silva, 2010, p. 440

Diferentemente de medidas que se referem à atribuição de um valor numérico para quantificar a intensidade ou duração da dor, avaliar dor é um processo complexo que valoriza a significância da dor para o indivíduo (STEVENS; GIBBINS, 2002; GUINSBURG et al., 2003; O’ROURKE, 2004). Porém, esta avaliação é influenciada por diversos fatores idiossincráticos como crenças pessoais, valores, educação e experiência, além de ser repleta de dificuldades quando se trata de lactentes pré-verbais (PILLAI RIDDELL; BADALI; CRAIG, 2004; PILLAI RIDDELL; CRAIG, 2006; PILLAI RIDDELL et al., 2008; ANAND; STEVENS; MCGRATH, 2007).

É possível constatar, através dos últimos anos, um aumento considerável no interesse pela utilização de métodos de aprendizagem de máquina com o intuito de analisar expressões faciais de dor em recém-nascidos, pois das várias modalidades de expressão não verbal (por exemplo, movimento corporal, vocalizações), a atividade facial pode fornecer a informação mais sensível, específica e acessível sobre a presença, natureza e gravidade da dor ao longo da vida, desde a infância (GRUNAU; CRAIG, 1987) até a idade avançada (HADJISTAVROPOULOS et al., 2014).

2.2 CLASSIFICAÇÃO DE DOR POR EXPRESSÕES FACIAIS

Há um interesse crescente em compreender as respostas comportamentais relacionadas à dor baseadas, principalmente, em expressões faciais. Por meio da inteligência artificial e visão computacional, novos trabalhos utilizando análise automática de dor em recém-nascidos têm surgido, principalmente pela carência de monitoramentos de dor contínuos e consistentes em clínicas e centros de saúde.

Na presente dissertação, o foco se encontra nas medidas comportamentais baseadas em análise de dor utilizando expressões faciais, que podem ser definidas como a tarefa de extrair automaticamente características relevantes de indicadores comportamentais de dor. Nesta seção, serão discutidos os métodos existentes que analisam expressões faciais para extrair características úteis para a classificação de dor em recém-nascidos.

Segundo Zamzmi et al. (2017), "expressões faciais de dor são definidas como movimentos e distorções nos músculos da face associados a estímulos de dor". Temos então uma relação entre dor em recém-nascidos e movimentos faciais, como aprofundamento do sulco nasolabial, rebaixamento das sobrancelhas, olhos estreitados, boca esticada vertical e horizontalmente, lábio apertado ou aberto, protrusão da língua, língua tensa e tremor no queixo (GRUNAU; CRAIG, 1987).

O arcabouço mais comum para reconhecimento automático de expressões de dor se divide em três etapas principais: detecção e normalização, extração de características e reconhecimento de padrões. A Tabela 2 destaca os principais trabalhos desenvolvidos até então, listando bases de dados, métodos de classificação e resultados gerados em cada um desses métodos.

Tabela 2 – Métodos automáticos para análise de expressões de dor em recém-nascidos

Referência	Base de Dados	Classificação	Resultados
(BRAHNAM et al., 2006)	COPE Indivíduo: 13 meninas e 13 meninos Raça: Caucásiano Faixa etária: 18 horas a 3 dias Estímulos: Estímulos de dor e 3 outros: ar, fricção e transporte de berço Dados: 200 imagens estáticas	PCA/LDA com L1 e SVM: Dor/não-dor, dor/repouso, dor/choro, dor/ar e dor/fricção Protocolo de Teste: Validação cruzada 10-fold	SVM acurácia média: Dor/não-dor (88%) dor/repouso (95%) dor/choro (80%) dor/sopro de ar (83%) dor/fricção (93%)
(BRAHNAM et al., 2007)	COPE	NNSOA, PCA/LDA e SVM: Dor (60 imagens) vs não-dor (144 imagens) Protocolo de Teste: Validação cruzada <i>leave-one-subject-out</i>	Acurácia média: NNSOA (90,20%) SVM (82,35%) PCA com L1 (80,35%) LDA com L1 (76,96%)
(GHOLAMI; HADDAD; TANNENBAUM, 2010)	COPE	RVM: Dor/não-dor Estimação de intensidade de Dor Protocolo de Teste: Validação cruzada <i>leave-one-subject-out</i>	Peso do coeficiente Kappa: 0.47 (especialista/RVM) 0.46 (não-especialista/RVM)
(NANNI; BRAHNAM; LUMINI, 2010)	COPE	SFFS e SVM Protocolo de Teste: Validação cruzada <i>leave-one-subject-out</i> (SFFS)	Maior área abaixo da curva ROC: 0,93
(MANSOR; REJAB, 2013b)	COPE	GANN e LDA Dor/não-dor em diferentes condições de iluminação	GANN acurácia média: 20 δ (91,02%) 40 δ (84,31%) 60 δ (80,91%) 80 δ (70,95%)
(HEIDERICH; LESLIE; GUINSBURG, 2015)	Gerada no Hospital Universitário da UNIFESP Indivíduo: 30 recém-nascidos Estímulos: Injeção intramuscular, retirada de sangue pelo calcanhar e punção venosa Dados: 360 imagens	Compara as distâncias entre os pontos nodais principais da imagem obtida no estado de repouso com as mesmas distâncias em imagens capturadas durante o monitoramento Protocolo de Teste: Validação através da concordância entre software e examinadores	Peso generalizado do coeficiente Kappa: 0,975 (examinadores/software)
(LU et al., 2016)	Gerada no Hospital Infantil de Nanjing Indivíduo: 76 recém-nascidos Estímulos: Imunização durante hospitalização e retirada de sangue pelo calcanhar Dados: 200 imagens para cada classe (Calmo, choro, dor moderada, dor severa)	SRC e SVM Protocolo de Teste: Validação cruzada 10-fold (RBF)	Melhor acurácia média: 85,50% (SRC e WLBP(5x3))
(ZAMZMI et al., 2018)	NPAD (<i>Infants Pain Assessment Database</i>) Gerada no Hospital Geral de Tampa Indivíduo: 16 meninas e 15 meninos Raça: 17% Caucásiano, 47% Branco, 17% Afro-americano, 12% Asiático e 7% outros Idade Gestacional Média: 36,4 Estímulos: Transporte de berço e retirada de sangue pelo calcanhar Dados: 200 imagens para cada classe (Calmo, choro, dor moderada, dor severa)	4 arquiteturas de CNN: VGG-F, VGG-M, VGG-S e VGG-Face	Melhor acurácia: 90,34% (VGG-Face)
(ZHI et al., 2018)	NPAD	SVM Protocolo de Teste: Validação cruzada <i>leave-one-subject-out</i>	<i>Decision Fusion:</i> 89,33% (três características) Melhor acurácia para visão de perfil: 87,88%

Fonte: Autor “adaptado de” Zamzmi et al., 2017, p. 4

Tomando-se novamente como referência o trabalho de revisão realizado por Zamzmi et al. (2017), os principais métodos propostos para extração de características relevantes de dor em imagens estão na Tabela 3. Nas subseções seguintes, esses resultados são descritos com detalhes.

Tabela 3 – Extração de características para análise de expressões de dor em recém-nascidos

Referência	Método de Extração
(BRAHNAM et al., 2006)	Vetorização das intensidades das imagens e redução de dimensionalidade (PCA)
(BRAHNAM et al., 2007)	Vetorização das intensidades das imagens e redução de dimensionalidade (PCA)
(GHOLAMI; HADDAD; TANNENBAUM, 2010)	Vetorização das intensidades das imagens
(NANNI; BRAHNAM; LUMINI, 2010)	Descritores LBP, LTP, ELTP e ELBP
(MANSOR; REJAB, 2013b)	GLCM
(HEIDERICH; LESLIE; GUINSBURG, 2015)	Distância entre os principais pontos fiduciais
(LU et al., 2016)	Descritor LBP
(ZAMZMI et al., 2018)	Extração profunda de características utilizando <i>transfer learning</i>
(ZHI et al., 2018)	Distância entre os principais pontos fiduciais

Fonte: Autor “adaptado de” Zamzmi et al., 2017, p. 8

2.2.1 Métodos Baseados em Redução de Características

Uma abordagem simples e prática para extração de características relevantes de dor em imagens estáticas é, primeiramente, transformar a imagem em uma matriz, tal que cada posição desta corresponde ao valor da intensidade do pixel. Esta matriz é transformada em um vetor, ou seja, todas as suas linhas originais são concatenadas, gerando ao final um vetor de tamanho $N_x \times N_y$, em que N_x e N_y são o comprimento e a altura da imagem em pixels, respectivamente. Após esta transformação, métodos de extração de características como Análise das Componentes Principais (PCA) e *Sequential Forward Floating Selection* (SFFS) podem ser aplicados para reduzir a dimensionalidade destes vetores (ZAMZMI et al., 2017).

SFFS é um método utilizado para seleção de características (VERVERIDIS; KOTROPOULOS, 2008) definido como uma família de algoritmos de busca gulosos, utilizados para redução de dimensão de espaços de características, sendo que a cada passo do algoritmo é adicionada ou removida uma única característica até que não haja melhoria no desempenho do classificador. O SFFS é uma extensão do SFS (Seleção Direta Sequencial), que também busca o melhor subconjunto de características, porém, desta vez, adicionando uma única característica a um subconjunto, inicialmente nulo, tal que esta satisfaça uma função de critérios. Segundo Zamzmi et al. (2017), "a diferença entre SFS e SFFS é que o SFFS permite, de acordo com a função de critério, excluir a pior característica do subconjunto (por exemplo, permite aumentar e diminuir dinamicamente o número de características até alcançar o melhor subconjunto)".

Um dos principais trabalhos, conhecido como projeto COPE (Classificação de Expressões de Dor), para avaliação de dor em recém-nascidos baseada na análise de expressões faciais foi apresentado em **b2005svm**. Para este trabalho, foram treinados e testados classificadores SVM utilizando imagens do banco de dados COPE, criado especificamente para este estudo, divididas em dois conjuntos: dor e não-dor; atingindo uma acurácia de classificação de aproximadamente 88%.

Este conjunto de dados consiste em 200 imagens coloridas capturadas de 26 bebês caucasianos, 50% meninas, usando uma câmera digital Nikon D100. A faixa etária dos bebês varia de dezoito horas a três dias e todos os bebês estavam em bom estado de saúde. As imagens faciais das crianças foram capturadas enquanto estas eram submetidas a quatro estímulos diferentes: estímulo da dor durante a punção do calcanhar (60 imagens), choro durante o transporte de uma criança de um berço para outro (18 imagens), estímulo de ar no nariz (23 imagens) e o estímulo de fricção, que envolve realizar atrito na superfície lateral externa do calcanhar com algodão embebido em álcool (36 imagens). Além desses quatro estímulos, o banco de dados inclui imagens dos bebês em estado neutro de descanso (63 imagens) (BRAHNAM et al., 2006).

A Figura 1 apresenta dois conjuntos de exemplos das cinco expressões: descanso, choro, estímulo de ar, fricção, e dor, existentes na base de dados COPE.

Figura 1 – Dois conjuntos de exemplos das cinco expressões presentes na base de dados COPE



Fonte: Brahnam et al., 2005

Para a extração de características relevantes de dor em (BRAHNAM et al., 2005) e (BRAHNAM et al., 2006), cada imagem foi rotacionada, cortada, convertida para escala de cinza, e reduzida para 100 x 120 pixels, e então transformada em um vetor de dimensão 12000. Realizando este mesmo procedimento para as 204 imagens, empilhando os vetores, construiu-

se uma matriz de dimensão 204 x 12000, contendo todas imagens do conjunto de dados. Por fim, aplicou-se o PCA para redução de dimensionalidade desta matriz.

Para classificação, métodos estatísticos baseados em distância, mais especificamente PCA, Análise Discriminante Linear (LDA) e Máquina de Vetores de Suporte (SVM), foram utilizados para classificar as imagens das faces dos bebês entre um dos seguintes pares: dor/não-dor, dor/descanso, dor/choro, dor/sopro de ar, dor/fricção. Os resultados mostraram que o método SVM, com função de núcleo polinomial de grau 3 avaliado por meio da validação-cruzada dividindo-se o conjunto original em dez subconjuntos (*k-fold*, com $k = 10$), alcançou a melhor taxa de reconhecimento e superou outros métodos baseados em distância, classificando dor versus não dor (88,00%), dor versus descanso (94,62%), dor versus choro (80,00%), dor versus sopro de ar (83,33%) e dor versus fricção (93,00%).

Brahnam, Nanni e Randall S Sexton (2008) estenderam os trabalhos anteriores (BRAH-NAM et al., 2005, 2006) empregando o algoritmo SFFS para extração de características e o NNSOA (*Neutral Network Simultaneous Optimization Algorithm*) para classificação além do PCA, LDA e SVM. Foi utilizada validação cruzada *leave-one-subject-out* ao invés do *k-fold* com $k = 10$. Os resultados mostraram que o NNSOA alcançou a maior taxa de classificação média (90,20%) em classificar imagens de bebês como dor (60 imagens) ou não dor (144 imagens). SVM, PCA, LDA com norma L^1 alcançaram as taxas de classificação média de 82,35%, 80,39% e 79,96%, respectivamente.

Ao contrário de detectar a presença ou ausência de expressões de dor, Gholami, Haddad e Tannenbaum (2009) apresentam um *kernel* de uma Máquina de Vetores de Relevância (RVM) para estimar o nível de intensidade das expressões de dor detectadas nas imagens da base de dados COPE, além da utilização do SVM para fins de comparação de resultados. O RVM nada mais é que um tratamento bayesiano de um modelo linear generalizado de forma funcional idêntica ao SVM. A acurácia encontrada para classificação de dor utilizando SVM com kernel linear foi de 90%. Aplicando o algoritmo RVM com kernel linear resultou em uma acurácia de classificação praticamente idêntica, isto é, 91%. Para medir a avaliação de intensidade de dor foi utilizado o coeficiente kappa. Para se confrontar a intensidade de dor indicada pelos examinadores (profissionais de saúde e não-profissionais) e o algoritmo RVM, comparou-se a nota, de 0 a 100, dada pelos examinadores com a incerteza para a classe “dor” (probabilidade *a posteriori*) dada pelo algoritmo RVM (GHOLAMI; HADDAD; TANNENBAUM, 2009), retornando 0,48 para profissionais e 0,52 para não-profissionais, o que mostra uma concordância moderada. Os resultados mostraram também uma acurácia praticamente idêntica para a classificação

binária. Em particular, os não-profissionais, profissionais e o RVM atingiram uma acurácia de 87%, 85% e 85%, respectivamente. Além disso, os resultados mostraram também que os profissionais e não-profissionais tendem a avaliar com a mesma intensidade de dor baseando-se nas expressões faciais presentes nas imagens.

Seguindo a mesma proposta de Gholami, Haddad e Tannenbaum (2009), em Gholami, Haddad e Tannenbaum (2010) alcançou-se um resultado ainda melhor, 90% de acurácia utilizando SVM e 91% utilizando RVM para a classificação binária. E, novamente, para a avaliação de intensidade de dor realizou-se a comparação utilizando a concordância entre a avaliação dos voluntários (profissionais de saúde e não profissionais) e o algoritmo RVM por meio do coeficiente de concordância kappa. Foram obtidas concordâncias de 0,47 para profissionais da saúde e 0,46 para não-profissionais.

As pesquisas citadas acima foram conduzidas utilizando a base de dados COPE que, na época, era a única base de imagens de face de dor de recém-nascidos disponível para acesso. Outros trabalhos que utilizaram a base de dados COPE com o mesmo intuito de analisar expressões de dor em recém-nascidos são apresentados em (BRAHNAM; NANNI; SEXTON, R., 2007; BRAHNAM; NANNI; SEXTON, R. S., 2008; MANSOR; REJAB, 2013a,b; NANNI; BRAHNAM; LUMINI, 2010).

Uma outra abordagem interessante e mais recente, diferente das citadas anteriormente, foi desenvolvida em Heiderich, Leslie e Guinsburg (2015), em que foi implementado um software capaz de avaliar dor procedural neonatal, o qual possui boa sensibilidade e especificidade para detectar movimentos faciais. A escala denominada como Sistema de Código Facial Neonatal (NFCS) (GRUNAU; CRAIG, 1987) foi selecionada como métrica básica de referência, em que 10 ações faciais são codificadas definindo a presença ou não de dor, sendo que cada ação facial recebe 1/0 (ocorreu/não ocorreu) durante o evento (GRUNAU et al., 1998). Baseado nestes conceitos, um software foi desenvolvido utilizando o ambiente Embarcado Delphi XE2 que, por meio da imagem facial, foram detectados 66 pontos nodais. A partir destes 66 pontos nodais identificados, foram selecionados 16 como sendo os principais, baseando-se em testes realizados em cinco recém-nascidos para identificar os pontos que exibiam os maiores movimentos quando os bebês sentiam dor. Entre estes 16 pontos, 14 distâncias foram escolhidas para avaliar semelhanças entre as imagens. Então, o software realiza a comparação entre essas distâncias na imagem obtida no estado de descanso, a calibrada, com as das imagens capturadas durante a sessão de monitoramento. Depois de calcular a similaridade, o software atribui pontos às ações faciais detectadas seguindo a escala NFCS (GRUNAU; CRAIG, 1987) para definir a

presença ou ausência de dor. A concordância entre a avaliação dos examinadores e o software foi excelente ($j = 0,975$). Além disso, o software exibiu 85% de sensibilidade e 100% de especificidade ao detectar expressões faciais neutras no estado de descanso e 100% de sensibilidade e especificidade ao detectar dor durante procedimentos dolorosos.

Uma questão em aberto presente em todos os trabalhos citados aqui é o tratamento, ou pré-processamento, de imagens de faces que não se encontram em posições frontais, ou próximas disso. A maioria dos sistemas de análise automática de expressões faciais focam em imagens de visão frontal da face, e poucos estudos investigam o efeito da visão de perfil nessas imagens (PANTIC; ROTHKRANTZ, 2004; PANTIC; PATRAS, 2006). De acordo com observações clínicas, movimentos com a cabeça ocorrem frequentemente durante experiências de dor e essa movimentação resulta em múltiplas visões da face, podendo acarretar em falhas na detecção da face e no reconhecimento de dor. Reconhecimento de expressões em imagens faciais de perfil é uma área desafiadora, pois muita informação representativa é perdida.

Visualizando esta deficiência presente nos trabalhos da área, Zhi et al. (2018) apresentam diferentes modelos de avaliação automática de dor em recém-nascidos, cada um deles atacando fatores que podem influenciar a avaliação final, como, por exemplo, as imagens faciais de perfil. Neste trabalho, representações dinâmicas de expressões faciais de dor foram propostas por meio da combinação de características de textura faciais temporais e características faciais geométricas temporais. Para a avaliação de dor em imagens de perfil, a melhor acurácia foi 87,88% para a hemiface esquerda, que fora igual à melhor taxa de reconhecimento da hemiface direita. Resultado consistente com a evidência de que expressões faciais espontâneas são mais simétricas (envolvendo a hemiface esquerda e direita) que as expressões deliberadas (HAGER, 1997).

2.2.2 Métodos Baseados em Variação de Padrões Binários Locais

Os trabalhos descritos brevemente nesta subseção utilizam o algoritmo de Padrões Binários Locais (LBP) (OJALA; PIETIKAINEN; MAENPAA, 2002) ou suas extensões em suas análises. Padrões Binários Locais (LBP) são um dos mais populares descritores de textura em visão computacional (ZAMZMI et al., 2017). Isto se dá pela sua simplicidade, baixa complexidade computacional, e robustez a variações de iluminação e erros de alinhamento (REN; JIANG; YUAN, 2016).

O LBP básico descreve a textura da imagem comparando o valor em escala de cinza de um pixel central X com os valores de seus P vizinhos pertencentes a um círculo de raio predefinido R , sendo a saída desta comparação um valor binário. Por exemplo, o valor de um pixel vizinho seria 1, caso o valor deste seja maior que o valor do pixel central, e 0, caso contrário. Esses valores binários são então combinados para formar padrões binários locais, convertidos para decimal e acumulados em um histograma discreto (ZAMZMI et al., 2017).

Uma extensão do LBP, conhecida como LBP melhorado, que é menos sensível a ruído é apresentado em Jin et al. (2004). Neste descritor, a redução de ruídos se dá pela comparação da intensidade dos pixels vizinhos com o valor médio local, e não pelo pixel central.

Tem-se também o Padrão Ternário Local (LTP) (TAN; TRIGGS, 2010), como uma outra extensão do LBP. A principal mudança comparando-se o LBP e o LTP se dá pelo fato de que a comparação entre o pixel central X e seus P vizinhos realizada no LBP é definida por um função de três entradas:

$$f(P,X,T) = \begin{cases} 1, & P \geq x + t \\ -1, & P \leq x - t \\ 0, & |X - P| < t \end{cases} \quad (1)$$

onde T é um *threshold* definido pelo usuário (ZAMZMI et al., 2017).

O Padrão Binário Alongado (ELBP) (LIAO; CHUNG, 2007) e o Padrão Ternário Alongado (ELTP) (NANNI; BRAHNAM; LUMINI, 2010) são variantes do LBP e LTP que utilizam a janela elíptica de vizinhança ao invés de uma circular. Segundo Liao e Chung (2007), a janela elíptica permite capturar a estrutura anisotrópica de imagens faciais com mais eficiência.

Nanni, Brahnam e Lumini (2010) e Nanni, Lumini e Brahnam (2010) utilizaram diversos descritores baseados em histogramas para detectar expressões faciais de dor em recém-nascidos, incluindo LBP, LTB, ELTP e ELBP. Estes descritores foram aplicados à base de dados COPE para extrair características relevantes de dor. Para a etapa de pré-processamento, as imagens foram redimensionadas, alinhadas, recortadas para se obter apenas a região facial, e divididas em blocos de dimensões 25 x 25. Foi então aplicado o algoritmo SFFS, descrito na subseção anterior, no conjunto de treinamento para selecionar os blocos mais discriminantes, avaliado por meio da validação cruzada *leave-one-out*. Para classificação de imagens de recém-nascidos em dor ou não-dor, um conjunto de SVMs de funções base radial foi construído e avaliado em um conjunto de testes. Os resultados mostraram que o descritor de textura ELTP ($P = 16$) alcançou a maior AUC, 0,926, em comparação com os outros descritores de textura. Esses

resultados mostraram ainda que sub-regiões da face são afetadas por expressões faciais. Então, dividir a imagem facial em blocos pequenos pode beneficiar o desempenho do algoritmo.

Lu et al. (2016) propuseram também um *framework* para detecção de expressões faciais de dor em imagens faciais de recém-nascidos através de variações de Padrões Binários Locais. O método proposto inclui os seguintes passos: pré-processamento das imagens, extração de características, redução de dimensionalidade e classificação. Para a etapa de pré-processamento, foi realizada a normalização em escala de cinza, a qual converte imagens coloridas em imagens em tons de cinza e realça o contraste e o brilho de regiões individuais. Para melhorar o contraste geral da imagem, adotou-se a equalização por histograma. No estágio de extração de características, descritores como Código de Gradiente Local (LGC), Padrão Direcional Local (LDP), Padrão de Textura Direcional Local (LDTP), além de várias variantes do LBP com diferentes tamanhos foram utilizados para extrair características de expressões das imagens faciais. O classificador baseado em representação esparsa (SRC) foi utilizado para classificar amostras de teste em quatro classes de expressões faciais: calmo, chorando, dor moderada e dor severa. Os resultados mostraram que a melhor acurácia média, 85,5%, é alcançada através da utilização do operador Padrão Binário Local Ponderado (WLBP) como descritor de características. Para avaliação do desempenho do algoritmo proposto, uma nova base de dados de imagens faciais de recém-nascidos foi construída, contendo 800 imagens faciais de 76 bebês para as quatro classes de expressões faciais (calmo, chorando, dor moderada e dor severa); 200 para cada classe. Todas as imagens foram recortadas e rotacionadas para que os olhos estivessem aproximadamente ao longo do mesmo eixo e foram redimensionadas para 112 x 92 pixels.

Uma limitação destes trabalhos supracitados (NANNI; BRAHNAM; LUMINI, 2010; NANNI; LUMINI; BRAHNAM, 2010; LU et al., 2016) é o uso de descritores de textura estáticos para detecção de expressões faciais. Estes lidam apenas com informações espaciais e ignoram o padrão dinâmico das expressões faciais. Segundo Zamzmi et al. (2017), "para mensurar as informações espaço-temporais, diversos descritores dinâmicos de textura como, por exemplo, Padrões Binários Locais em Três Planos Ortogonais (LBP-TOP) (ZHAO; PIETIKAINEN, 2007) podem ser explorados". Na verdade, existe uma limitação em utilizar imagens 2D estáticas para reconhecimento de expressões de dor. Imagens estáticas ignoram informações temporais e a dinâmica das expressões. Isso afeta a capacidade de compreender a expressão facial e sua evolução ao longo do tempo. Oclusões tais como auto-oclusão, máscara de oxigênio, chupeta, entre outros, são comuns nestes ambientes clínicos, e seguem como artefatos restritivos para análise de imagens estáticas (ZAMZMI et al., 2017).

2.2.3 Aprendizado Profundo

Mais recentemente, Zamzmi et al. (2018) propuseram um novo arcabouço para reconhecimento de expressões de dor em recém-nascidos, utilizando transferência de aprendizagem através de Redes Neurais Convolucionais (CNN) pré-treinadas. Especificamente, foi proposto explorar uma CNN treinada originalmente em uma base de dados relativamente similar para reconhecimento de faces (*VGG-Face*), assim como CNN's pré-treinadas em uma base de dados relativamente diferente para classificação de imagem (*VGG-F,M,S*) para extrair características profundas das faces dos recém-nascidos. No estágio final, vários classificadores supervisionados de aprendizagem de máquina foram treinados para classificar expressões faciais em recém-nascidos entre dor ou não-dor.

O arcabouço proposto alcançou uma acurácia de 90,34% e *Area Under The Curve* (AUC) igual à 0,841 em uma base de dados construída pela equipe de neonatologia do Hospital Geral de Tampa. Foi também proposta a combinação de características profundas e tradicionais, admitindo a hipótese de que características mistas melhorariam o desempenho da classificação de dor. Através desta proposta foi alcançada a acurácia de 92,71% e AUC igual a 0,948.

2.3 RECONHECIMENTO DE DOR EM IMAGENS FACIAIS

Podemos dividir reconhecimento de dor em imagens faciais em duas tarefas principais de classificação: detecção de dor e estimativa de intensidade de dor. Apresentamos a seguir, baseando-se no trabalho de revisão realizado em Zamzmi et al. (2017), uma descrição sucinta das limitações dessas tarefas.

2.3.1 Detecção de dor

Detecção de dor busca identificar a presença ou ausência de dor. Logo, tem-se como saída de um classificador treinado com características relevantes de dor, um valor binário, indicando se o recém-nascido está ou não com dor.

Como apresentado neste capítulo, foram construídos diversos classificadores com o intuito de detectar dor em imagens faciais de recém-nascidos. Tais classificadores conseguiram atingir variados níveis de desempenho em detectar a classe dor, que chegam a valores máximos de aproximadamente 92% de acurácia.

Detecção de dor nos oferece a classe dor, porém não existe o retorno da intensidade ou nível da dor detectada. Segundo Zamzmi et al. (2017), "para aplicações de avaliação de dor, detectar a dor sem sua intensidade pode não ser suficiente dadas três razões principais". Primeiro, não tem-se a gravidade da dor e as diferenças individuais em resposta aos estímulos dolorosos para esses casos. Segundo, classificar dor sem suas intensidades não fornece informação sobre a quantidade de dor e como essa se altera com o tempo; um recém-nascido pode experimentar diferentes intensidades de dor em diferentes intervalos de tempo (ZAMZMI et al., 2017). Por isto, podemos considerar que estimar a intensidade de dor é importante e pode nos conduzir a melhores compreensões e intervenções.

2.3.2 Estimativa de intensidade de dor

Estimar a intensidade do estímulo doloroso oferece uma melhor e mais completa avaliação da dor e pode resultar, conseqüentemente, em uma melhor administração da própria dor (ZAMZMI et al., 2017).

Métodos de reconhecimento de dor severa foram propostos para se estimar a intensidade da dor e que podem ser aplicados em problemas envolvendo recém-nascidos. Como apresentado anteriormente, Gholami, Haddad e Tannenbaum (2010) apresentaram um método utilizando RVM, tal que o classificador RVM oferece como saída as probabilidades dos membros das classes ou *labels*. A incerteza presente nestas probabilidades foi então utilizada para estimar a intensidade da dor nos recém-nascidos.

Da mesma forma, Gruss et al. (2015) introduziram um método para estimar quatro níveis de dor utilizando SVM. Expressão facial e sinais biopotenciais (eletromiograma, nível da condutância da pele e eletrocardiograma) foram registrados sob quatro níveis de dor. Em seguida, após a extração de características, classificadores SVM foram treinados com 75% dos dados e testados com os 25% restantes. O método proposto alcançou diferentes níveis de acurácia para os quatro níveis de dor analisados (todos acima de 75%).

2.4 BASES DE DADOS

A qualidade, complexidade e capacidade são três fatores importantes que devem ser considerados na coleta de bancos de dados para avaliação de dor em recém-nascidos. Bases de baixa qualidade com uma noção vaga de sofrimento e anotações inadequadas podem levar

a resultados imprecisos. Além disso, a complexidade do banco de dados, em termos de suas modalidades/dimensões, é fundamental para desenvolver um sistema confiável de avaliação de dor capaz de avaliar a dor em caso de falta de dados. Bancos de dados com número relativamente pequeno de indivíduos não são suficientes para avaliar o desempenho do sistema e tirar conclusões (ZAMZMI et al., 2017). Portanto, coletar imagens de alta qualidade e em grandes quantidades se faz necessário para desenvolver sistemas robustos de avaliação de dor em recém-nascidos.

A maioria dos bancos de dados de dor existentes não está publicamente disponível, por razões legais/éticas, e essas bases foram desenvolvidas apenas para pesquisa. Porém, alguns trabalhos possibilitam acesso a suas bases para pesquisadores por meio de um contrato de licença que garante todos os direitos reservados aos indivíduos registrados nas bases, além de outros acordos definidos pelos desenvolvedores do banco de dados em particular, como são os exemplo das bases COPE (BRAHNAM et al., 2006) e NPAD (Neonatal Pain Assessment Dataset) (ZAMZMI et al., 2018).

Esta seção fornece breves descrições dos bancos de dados de dor em recém-nascidos disponíveis publicamente, além da base desenvolvida pela UNIFESP, alvo deste trabalho.

2.4.1 COPE

Esta base de dados cumpriu os protocolos e diretrizes éticas para pesquisa envolvendo seres humanos no St. John's Health System Inc. O consentimento informado foi obtido para um dos pais, geralmente a mãe em consulta com o pai. A maioria dos pais foi recrutada na unidade neonatal do Hospital St. John's alguns dias após o parto. Apenas mães que experienciaram partos sem complicações foram abordadas (BRAHNAM et al., 2006).

Um total de 200 fotografias coloridas foram tiradas de 26 recém-nascidos caucasianos (13 meninos e 13 meninas) com idade variando entre 18 horas e 3 dias de idade. Seis meninos foram circuncidados um dia antes das fotografias serem tiradas, e o último tempo de alimentação antes da sessão fotográfica variou de 45 minutos a 5 horas. Todos os bebês estavam com boa saúde.

Todas as fotografias foram tiradas usando uma câmera digital Nikon D100 sob condições de iluminação ambiente em uma sala separada dos outros recém-nascidos.

As imagens faciais das crianças foram capturadas enquanto estas eram submetidas a quatro estímulos diferentes: transporte de uma criança de um berço para outro, estímulo de ar

no nariz, atrito na superfície lateral externa do calcanhar com algodão embebido em álcool e a punção do calcanhar. O estado do bebê depois de ser transportado para outro berço foi posteriormente avaliado no momento em que as fotografias foram tiradas em um dos dois estados: choro ou descanso.

Todos os estímulos foram administrados por uma enfermeira assistente. Seguindo o exemplo de Balda et al. (2000) e os requisitos padrões dos procedimentos médicos, as fotografias das expressões faciais dos quatro estímulos foram obtidas na seguinte sequência (BRAH-NAM et al., 2006):

- a) transporte de um berço para outro (descanso / choro): depois de ser transportado de um berço para outro, o recém-nascido foi coberto, mantendo a face visível, e uma série de fotografias foi tirada ao longo de 1 minuto. O estado do recém-nascido foi anotado como chorando ou descansando para cada fotografia tirada nessas séries;
- b) estímulo de ar: depois de descansar por pelo menos 1 minuto, o nariz do neonato foi exposto a uma nuvem de ar emitido por um limpador de lentes de plástico maleável. Uma série de fotos do rosto do recém-nascido foi tirada imediatamente após o sopro de ar entrar em contato com o rosto do bebê;
- c) atrito: após descansar por pelo menos 1 minuto, o recém-nascido recebeu atrito na superfície lateral externa do calcanhar com algodão embebido em álcool 70 por 10 a 15 segundos. O rosto do recém-nascido foi repetidamente fotografado durante as fricções do atrito;
- d) dor: após repouso por pelo menos 1 minuto, a superfície lateral externa do calcanhar foi puncionada para coleta de sangue. Várias fotografias contínuas da face do recém-nascido foram tiradas imediatamente após a introdução da agulha e enquanto a pele do calcanhar foi comprimida para amostras de sangue.

Das 200 fotografias faciais, 63 são de repouso, 18 de choro, 23 de estímulos de ar, 36 de atrito e 60 de dor. A Figura 6 apresenta um conjunto de exemplos dessa base COPE.

Figura 2 – Exemplos da base de faces COPE



Fonte: Autor

2.4.2 NPAD

Este conjunto de dados piloto (NPAD-1) faz parte de um projeto de avaliação de dor neonatal que visa desenvolver novos algoritmos e técnicas para o monitoramento automático da dor neonatal. Os pesquisadores e desenvolvedores deste projeto estão atualmente envolvidos em um esforço para coletar um grande conjunto de dados de dor multimodal de recém-nascidos e prematuros, NPAD-2, que também estará disponível em breve por requisição. O conjunto de dados NPAD-1 contém 12 vídeos (187096 quadros) de 7 indivíduos registrados na Unidade de Terapia Intensiva Neonatal (NICU) do Hospital Geral de Tampa.

Ao todo, 31 recém-nascidos (16 do sexo feminino e 15 do sexo masculino) foram gravados sobre um breve estímulo agudo, como coleta de sangue ou imunização durante a internação no hospital. A idade gestacional média dos bebês é de 36,4 semanas, variando de 30,4 a 40,6 (Desvio Padrão = 2,7). A distribuição étnica é de 17% caucasianos, 47% brancos, 17% afro-americanos, 12% asiáticos e 7% outros. Qualquer criança nascida na faixa de 28 e 41 semanas de gestação foi elegível para inscrição após obter o consentimento informado dos pais. Foram excluídos bebês com anormalidades faciais cranianas e distúrbios neuromusculares (ZAMZMI et al., 2018; SALEKIN, 2017).

Foi utilizada a câmera GoPro Hero3+ para registrar a expressão facial, o movimento do corpo e o som de choro dos bebês. Todas as gravações foram realizadas no ambiente clínico normal, modificado apenas pela adição da câmera.

Foi registrado cada lactente em sete períodos de tempo: 1) Antes do procedimento doloroso para obter a observação do estado inicial; 2) Período de preparação do procedimento que começa com o primeiro toque, pode incluir o posicionamento ou a preparação da pele e termina com a perfuração da pele; 3) Procedimento doloroso, inclui todo o processo; 4) Um minuto após a conclusão do procedimento doloroso; 5) Dois minutos após a conclusão; 6) Três minutos após a conclusão; e 7) Período de recuperação cinco minutos após o procedimento. Cada período de tempo foi observado por enfermeiros treinados para fornecer a avaliação da dor usando a escala NIPS (Neonatal Infant Pain Scale) (LAWRENCE et al., 1993).

Essa escala NIPS consiste em seis elementos, que são expressão facial, choro, movimento do corpo (ou seja, braços e pernas), estado de excitação e padrões de respiração. Cada elemento do NIPS foi pontuado em uma escala de 0-1, com exceção do choro, que é pontuado em uma escala de 0-1-2. Uma pontuação total de 3-4 representa dor moderada e uma pontuação maior que 4 indica dor intensa. Para obter o *ground truth* para cada quadro de vídeo, foi utilizado o limiar da pontuação total (ou seja, dor severa, dor moderada ou ausência de dor) como o rótulo para avaliação de algoritmo. Em Zamzmi et al. (2018), incluiu-se apenas os rótulos de dor / ausência de dor porque o número de quadros para dor moderada no conjunto de dados atual é pequeno.

Foi então aplicado o detector de face em 200 vídeos para encontrar o rosto e os pontos de referência na imagem. Em seguida, foram selecionados alguns quadros de cada vídeo, removendo desse modo uma grande quantidade de quadros semelhantes. Os quadros selecionados de todos os vídeos (isto é, 3026) foram redimensionados para 224×224 .

A Figura 3 apresenta um conjunto de exemplos da base NPAD-1.

Figura 3 – Exemplos da base de dados NPAD-1



Fonte:Autor “adaptado de” Salekin, 2017

2.4.3 UNIFESP

O banco de imagens da UNIFESP (HEIDERICH; LESLIE; GUINSBURG, 2015) foi desenvolvido para construção de um software capaz de identificar automaticamente a expressão de dor em recém-nascidos por meio de imagens bi-dimensionais capturadas por um aparato de três câmeras, apresentado na Figura 4.

Figura 4 – Posicionamento das três câmeras envolta do berço



Fonte: Heiderich, Leslie e Guinsburg, 2015

O software desenvolvido utiliza a distância entre pontos específicos identificados no rosto do recém-nascido e baseado na escala unidimensional NFCS (GRUNAU; CRAIG, 1987), classificando a existência ou não de dor no recém-nascido.

A base foi construída a partir de fotos capturadas antes, durante e depois de procedimento aplicado aos recém-nascidos, podendo esses procedimentos serem uma punção venosa, capilar ou injeção intramuscular, procedimentos comuns e necessários, porém, considerando a fragilidade do recém-nascido, são procedimentos invasivos, dolorosos e/ou estressantes. A captura dessas fotos foi feita a cada 3 segundos (HEIDERICH; LESLIE; GUINSBURG, 2015).

A partir dos vídeos gravados de cada recém-nascido, foram selecionadas 360 amostras de imagens de 30 recém-nascidos entre 34 e 41 semanas de idade gestacional e entre 24 e 168 horas de vida (prematuros tardios ou a termo), alimentados e saudáveis (sem más-formações congênitas, anomalias faciais, necessidade de suporte ventilatório, sonda gástrica ou injeções intramusculares e/ou subcutâneas) (HEIDERICH; LESLIE; GUINSBURG, 2015). Esses recém-nascidos estavam internados no Hospital São Paulo (Hospital Universitário da Escola Paulista de Medicina da Universidade Federal de São Paulo). Todos os recém-nascidos selecionados participaram da pesquisa com consentimento dos familiares ou responsáveis (HEIDERICH; LESLIE; GUINSBURG, 2015) com a permissão de uso das imagens pelo comitê de ética, sob o número de rastreamento CCAE: 66743417.0.0000.5505 e o número de parecer 2.035.113. Não houve restrição de seleção dos recém-nascidos quanto a gênero, raça ou cor.

Entre as imagens capturadas pelo sistema, foram escolhidas também 12 imagens faciais de cada um dos 30 recém-nascidos totalizando 360 imagens separadas, em que 138 foram adquiridas antes do procedimento, 30 durante o procedimento e 192 até dez minutos após o procedimento. Em formato de relatório e em disposição aleatória essas imagens foram então submetidas para avaliação de profissionais da área de saúde com especialização em neonatologia. O formato das imagens apresentadas no relatório, assim como um exemplo preenchido pelos profissionais e também utilizado nos resultados do sistema, podem ser vistos na Figura 5.

Figura 5 – Exemplo de imagem do relatório (esquerda) e da imagem avaliada no relatório (direita) utilizada em Heiderich, Leslie e Guinsburg (2015)



Fonte: Heiderich, Leslie e Guinsburg, 2015

A avaliação feita pelos profissionais foi baseada na escala NFCS, identificando em cada imagem: frente saliente; fenda palpebral; sulco nasolabial; boca aberta e boca estirada. A cada três ou mais características positivas a imagem do recém-nascido era então classificada "com dor", caso contrário "sem dor". Essa avaliação agregou valor de informação nas imagens, classificando 164 imagens "com dor" e 196 "sem dor", sendo essa classificação humana tomada como referência e usada na construção do algoritmo sequencial de procedimentos computacionais para determinação da existência de dor, de maneira automática. No relatório, as imagens classificadas "Com Desconforto" foram consideradas "Com Dor", devido a características de comparação com a escala unidimensional NFCS.

Ao final, baseadas na escala unidimensional NFCS, as imagens do relatório foram analisadas pelos profissionais treinados, resultando na seguinte classificação: 196 imagens "Sem Dor" e 164 imagens "Com Dor". Já o sistema classificou: 128 imagens "Sem Dor" e 264 imagens "Com Dor". Relacionando as imagens e suas classificações geradas pelos profissionais e

pelo sistema, 232 classificações exibiram concordância, sendo 150 imagens classificadas “Com Dor” e 82 imagens “Sem Dor” (HEIDERICH; LESLIE; GUINSBURG, 2015).

A Figura 6 apresenta um conjunto de exemplos ¹ da base de dados UNIFESP.

Figura 6 – Exemplos da base de dados UNIFESP



Fonte: Autor

2.5 CONSIDERAÇÕES COMPLEMENTARES

Segundo Lynch (2001), "a atual atenção para melhores métodos de avaliação da dor em lactentes pré-verbais contribui para aumentar a sensibilidade dos profissionais de saúde em relação à natureza das experiências dolorosas".

A criação de modelos capazes de auxiliar profissionais da saúde na avaliação de dor em recém-nascidos se faz necessária dada a importância do treinamento formal destes, no que se refere às questões relativas à dor no período neonatal, em todos os níveis de formação, e a adoção de rotinas diagnósticas e terapêuticas nas unidades responsáveis por cuidarem de recém nascidos (CHERMONT et al., 2003). Essa dissertação segue as etapas principais de pré-processamento, análise e classificação de imagens faciais de recém-nascidos, discutidas brevemente neste capítulo, com atenção especial para a etapa de normalização dos dados como um passo essencial dentro do arcabouço, pois é necessário que seja garantida a correspondência entre as regiões de interesse da face no momento da classificação.

¹ O retângulo de reconhecimento facial marcado nas imagens pode ser considerado como um artefato, mas não foi aplicada nenhuma técnica para retirar o mesmo, a fim de preservar as características originais avaliadas pelos profissionais da área de saúde.

3 METODOLOGIA

Este capítulo descreve em detalhes a metodologia proposta e implementada nesta dissertação, composta das seguintes etapas principais: normalização espacial das imagens, extração de características e análise estatística multivariada.

Para a normalização espacial das imagens, implementou-se um *framework* computacional que combina posicionamento de pontos de referências (*landmarks*) faciais, registro não-rígido de imagens, utilizando funções triangulares, e construção de atlas inéditos de faces de recém-nascidos (GIBSON, 2006; XAVIER et al., 2015, 2016b; RUECKERT et al., 1999; SERAG et al., 2012; ALJABAR et al., 2008; SHAFFER et al., 2016). Para extração de características, utilizou-se o PCA (Análise das Componentes Principais), para redução de dimensionalidade e interpretação do conjunto de dados (JOHNSON; WICHERN, 2006c). Foi possível, então, identificar as componentes cujos dados apresentam maior variância em relação aos conjuntos de imagens, sendo cada imagem definida como “Dor” ou “Não-Dor”. Após a redução de dimensionalidade, aplicou-se o MLDA (*Multi-Linear Discriminant Analysis*), com o intuito de encontrar as características mais discriminantes e, por fim, classificar as imagens em escala contínua entre “Dor” e “Não-Dor”.

3.1 NORMALIZAÇÃO ESPACIAL DAS IMAGENS

Diferenças de iluminação e também variações posicionais dos bebês no momento da captura das imagens não são obviamente discriminantes para o fenômeno dor, e acabam atrapalhando a análise computacional posterior. Isto dado o fato de que a maioria dos trabalhos nesta área foi desenvolvida utilizando modelos de face que se baseiam em diferença da distância entre pontos fiduciais da face, ou seja, é necessário que os dados originais sejam normalizados espacialmente, garantindo que haja uma correspondência no momento da comparação entre esses pontos.

3.1.1 Pré-processamento

A primeira etapa de normalização é o pré-processamento das imagens das bases de interesse: COPE e UNIFESP. Esta etapa consiste em remover as características que dificultam a

realização das próximas etapas, como orientação das imagens, dimensões diferentes e artefatos na imagem.

Para a base COPE, o primeiro pré-processamento realizado foi um recorte das imagens originais, mais precisamente uma redução na quantidade de *pixels* de 3000×2000 para 752×500 (quatro vezes menor), com o objetivo de otimizar o consumo de memória e tempo de processamento das etapas subsequentes. Em seguida, foi observado que as imagens de face não estavam orientadas verticalmente, o que degrada a qualidade da predição dos pontos fiduciais da face. Portanto, as imagens foram rotacionadas manualmente em até 90° ou -90° , dependendo da orientação da imagem.

Para a base UNIFESP, 130 *pixels* na borda direita de todas as imagens são brancos (valor 255 na representação de 8 bits), o que permite que esta informação da imagem possa ser removida, resultando em imagens com 320 *pixels* de largura. Além disso, as imagens da base UNIFESP não apresentam uniformidade na quantidade de *pixels* na vertical (algumas contêm 233 e outras 221 *pixels*). Portanto, as imagens com 233 *pixels* de altura foram reduzidas para 221 *pixels*. O tamanho final das imagens da UNIFESP pré-processadas ficou padronizado em 320×221 .

3.1.2 Predição dos pontos fiduciais da face

Neste trabalho, o alinhamento das imagens é realizado exclusivamente pelas informações geométricas das imagens, ou seja, a informação de textura não é utilizada. Desta forma, a informação utilizada são os pontos fiduciais da face.

A predição desses pontos é feita por um algoritmo baseado em *deep learning*, disponível publicamente¹ (BULAT; TZIMIROPOULOS, 2017), que por sua vez utiliza uma rede convolucional para detecção de faces (ZHANG et al., 2017), e em seguida marca 68 pontos fiduciais, que podem ser visualizados pelos *pixels* brancos da Figura 7a.

3.1.3 Alinhamento global

A transformação utilizada para alinhamento global é também chamada de transformação afim por similaridade, que permite translação, rotação e escala e tem a propriedade de manter

¹ <https://github.com/ladrianb/face-alignment>

as proporções da imagem. Esta é uma transformação linear, ou seja, a mesma transformação é aplicada sobre todos os pixels. A transformação é definida por uma matriz:

$$\mathbf{T} = \begin{bmatrix} \Lambda \cdot \cos(\theta) & \Lambda \cdot \sin(\theta) & t_x \\ -\Lambda \cdot \sin(\theta) & \Lambda \cdot \cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

tal que Λ é o fator de escala, θ o ângulo de rotação, e t_x e t_y são as translações horizontais e verticais, respectivamente.

Dos 68 pontos fiduciais, 5 pontos foram escolhidos para o alinhamento global. Estes cinco pontos representam regiões de interesse importantes da face, como centro dos olhos, nariz e cantos da boca.

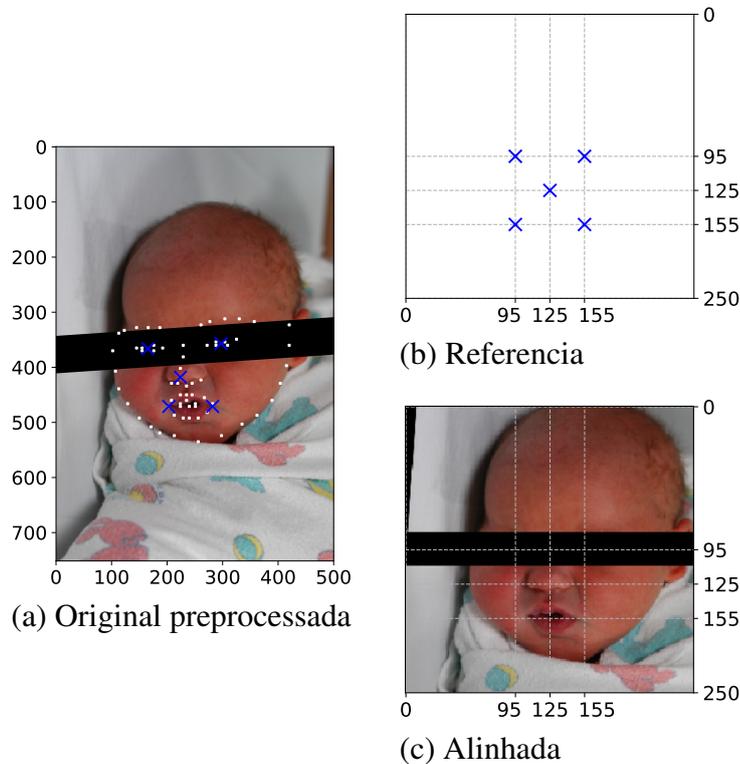
Considerando \mathbf{u} as coordenadas x e y dos 5 pontos iniciais (pontos de saída) e \mathbf{u}' as coordenadas x e y dos 5 pontos da imagem alinhada globalmente (pontos de destino), é possível descrever a equação:

$$\mathbf{u}' = \mathbf{T} \cdot \mathbf{u} \quad (3)$$

em que os parâmetros da matriz \mathbf{T} podem ser estimados ao formular esta equação como um problema de mínimos quadrados (UMEYAMA, 1991).

As etapas do alinhamento global podem ser visualizadas na Figura 7.

Figura 7 – Etapas para normalização global das imagens: (a) marcação automática dos 68 pontos fiduciais na imagem original preprocessada, (b) especificação das 5 coordenadas de destino (olhos, nariz e cantos da boca) e (c) a imagem normalizada globalmente



Fonte: Autor

3.1.4 Alinhamento não-rígido

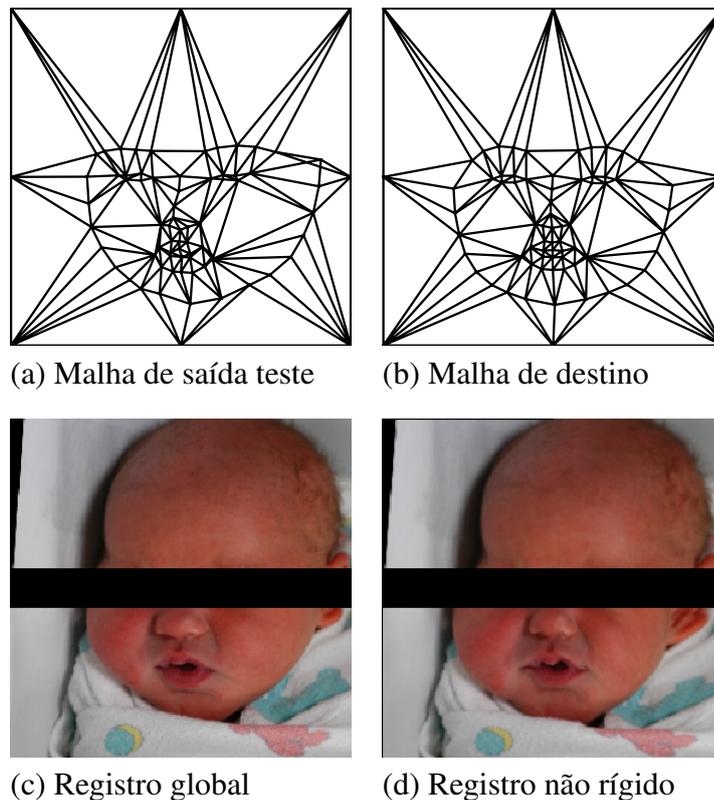
O método de alinhamento não-rígido adotado aqui é baseado em malhas triangulares², que já foi utilizado para alinhamento de faces em artigos anteriores de reconhecimento facial (TAIGMAN et al., 2014).

Com as imagens globalmente alinhadas, é possível encontrar os 68 pontos fiduciais médios para cada base de dados, que por sua vez serão utilizados para calcular uma malha triangular por meio do algoritmo de triangulação de Delaunay (BARBER; DOBKIN; HUHDANPAA, 1996). A malha triangular formada pelos 68 pontos fiduciais é chamada neste trabalho de malha de destino. Em seguida, para os 68 pontos fiduciais de cada imagem, computa-se uma malha triangular com base nos vértices da malha de destino para garantir que a quantidade de triângulos é a mesma. Esta malha específica para cada imagem é chamada neste trabalho de malha

² Conhecido em inglês por *piecewise*.

de saída. Para cada triângulo da malha de saída foi calculada uma transformação afim capaz de deformar exatamente para a forma do triângulo correspondente à malha de destino. A transformação afim de cada triângulo é calculada de forma similar ao método descrito na seção anterior, com a diferença de que a estimação dos parâmetros da transformação afim é feita com base em 3 pontos, que correspondem aos vértices de cada triângulo, portanto existe uma solução única para deformação de cada triângulo (UMEYAMA, 1991). Adicionalmente, para manter a informação fora da região formada pelos 68 pontos fiduciais, foram acrescentados 8 pontos nas extremidades de cada imagem, resultando, por exemplo, nas imagens da Figura 8.

Figura 8 – Etapas para normalização não-rígida das imagens: (a) Malha triangular de saída, (b) especificação da malha triangular de destino, (c) a imagem correspondente à malha de saída e (d) a imagem resultante normalizada não rigidamente.



Fonte: Autor

3.2 NORMALIZAÇÃO TODOS CONTRA TODOS

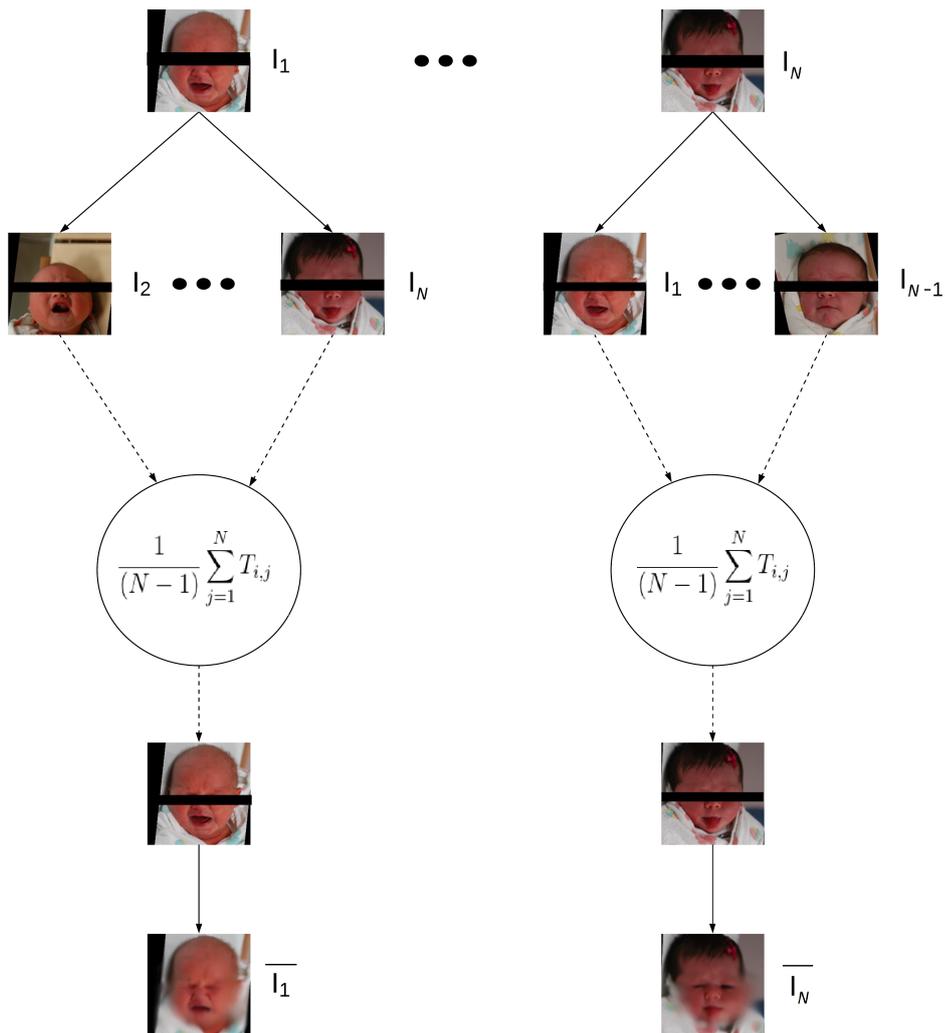
Para servir de entrada aos métodos de extração multivariada de informação, foram geradas novas imagens através da mesma metodologia utilizada na seção anterior, porém desta vez aplicando a política de todos-contra-todos, em vez de um-contra-todos. Para o um-contra-todos,

utiliza-se uma imagem fixa de referência e todos as normalizações utilizam esta imagem como base, o que o todos-contra-todos propõe é não utilizar mais uma única imagem como referência.

Este método iterativo propõe que para cada passo, uma imagem, digamos I_1 , da base de dados de tamanho N seja utilizada como referência, ou seja, é retirado do conjunto uma imagem, e esta é utilizada como referência para calcular a transformação das $N - 1$ imagens restantes. Após isto, calcula-se a média dessas transformações e aplica-se na imagem I_1 . Este processo se repete então por mais $N - 1$ vezes, até que todas as imagens estejam normalizadas.

Foi escolhido este método para normalização dos dados principalmente pelo fato de que este busca eliminar o viés que uma imagem fixa de referência pode, possivelmente, criar em uma normalização um-contra-todos. A Figura 9 ilustra esse processo, sendo que I_1, \dots, I_N representam o conjunto de imagens da base de dados.

Figura 9 – Processo de normalização utilizando todos-contra-todos



Fonte: Autor

Para cada imagem I_i escolhida como referência, o processo ilustrado na Figura 9 gera transformações $T_{i,j}$ para $j = 1, \dots, N, j \neq i$, que por fim geram uma transformação média \bar{T}_i :

$$\bar{T}_i = \frac{1}{(N-1)} \sum_{j=1}^N T_{i,j} \quad (4)$$

para cada imagem i . A imagem média \bar{I}_i é definida como a imagem I_i espacialmente transformada por \bar{T}_i .

3.3 EXTRAÇÃO MULTIVARIADA DE INFORMAÇÃO

Para os casos em que o conjunto de dados a ser analisado inclui medições simultâneas em muitas variáveis, a metodologia empregada recebe o nome de análise multivariada (JOHNSON; WICHERN, 2006a). Este tipo de análise foi aplicada no presente trabalho com o objetivo de realizar análises conjuntas e simultâneas de duas ou mais variáveis correlacionadas entre si, provenientes de um mesmo conjunto de dados. A necessidade de compreender as relações entre múltiplas variáveis torna a análise multivariada um assunto complexo (JOHNSON; WICHERN, 2006a), porém por meio da mesma é possível interpretar e quantificar qual tipo de relação existe entre essas variáveis e qual a relevância dessas relações com o problema em estudo.

As subseções seguintes descrevem as principais técnicas utilizadas no presente trabalho, PCA (*Principal Component Analysis*), LDA (*Linear Discriminant Analysis*) e MLDA (*Multi-Linear Discriminant Analysis*).

3.3.1 Análise das Componentes Principais (PCA)

Uma análise das componentes principais procura explicar a estrutura de variância-covariância de um conjunto de variáveis por meio de combinações lineares dessas variáveis. Seus principais objetivos são: (1) redução de dimensionalidade e (2) interpretação dos dados.

Embora hajam n variáveis para se reproduzir completamente a variabilidade total do sistema, geralmente podemos explicar esse mesmo conjunto de dados com apenas p componentes principais, tal que p é um número inteiro menor que n . Caso isso seja realmente possível, existe, aproximadamente, tanta informação nas p componentes principais quanto nas n variáveis originais. Logo, as p componentes principais podem substituir as n variáveis iniciais, e o conjunto

de dados original, consistindo de N amostras em n variáveis, é reduzido para um conjunto de N amostras em p componentes principais.

Esta análise geralmente revela relações que não foram previamente supostas e, desse modo, permite interpretações que normalmente não seriam possíveis sem este tipo de análise.

A partir da vetorização das imagens, obtém-se uma matriz \mathbf{X} do tipo $N \times n$, em que N é o número de exemplos da base e n a quantidade de pixels da imagem (*altura* \times *largura*). O PCA basicamente calcula uma matriz de transformação ϕ_{PCA} dada pelos autovetores da matriz de covariância de \mathbf{X} , sendo que esta matriz de transformação foi composta selecionando os p autovetores com autovalores não-nulos, tal que $p \leq \min(N, n)$. Os dados são então projetados nos $[\phi_1, \phi_2, \dots, \phi_p]$ autovetores rearranjados em ordem decrescente pelos seus autovalores correspondentes, ou seja $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, de acordo com a direção de maior variância de \mathbf{X} .

Cada hiperplano ϕ_q , em que $q = 1, 2, \dots, p$, pode ser utilizado para descrever um mapa espacial das características da face dos recém-nascidos que mais variam nos dados, movendo de um lado do eixo da componente principal para o outro (COOTES; EDWARDS; TAYLOR, 2001; DAVATZIKOS, 2004). Portanto, é possível navegar ao longo de tais dimensões para capturar e compreender as mudanças mais expressivas na matriz de dados \mathbf{X} . Este método de navegação no hiperplano do PCA (COOTES; EDWARDS; TAYLOR, 2001; SATO et al., 2008; XAVIER et al., 2015; GREGORI; SANCHES; THOMAZ, 2016) pode ser matematicamente descrito como:

$$\mathbf{y}_{q,i} = \bar{\mathbf{x}} + (i\sqrt{\lambda_q}) \times \phi_q \quad (5)$$

onde $i \in -3, -2, -1, 0, 1, 2, 3$, q é o índice da componente principal correspondente para a navegação, e $\bar{\mathbf{x}}$ é o vetor de média global n -dimensional de \mathbf{X} .

3.3.2 Análise Linear Discriminante (LDA)

Com a Análise Linear Discriminante (LDA) é possível encontrar direções discriminantes entre dois grupos de interesse (SATO et al., 2008), neste caso, dor e não-dor ($g = 2$). A ideia aqui é sair de um espaço p -dimensional em um espaço $(g - 1)$ -dimensional

Formalmente, o LDA procura separar a distribuição de grupos de amostras por meio da maximização da separabilidade inter-classe e da minimização da variabilidade intra-classe, tal que \mathbf{S}_b é a matriz de covariância inter-classe definida por:

$$\mathbf{S}_b = \sum_{i=1}^2 N_i (\bar{\mathbf{z}}_i - \bar{\mathbf{z}})(\bar{\mathbf{z}}_i - \bar{\mathbf{z}})^T, \quad (6)$$

e \mathbf{S}_w é a matriz de covariância intra-classe definida como:

$$\mathbf{S}_w = \sum_{i=1}^2 \sum_{j=1}^{N_i} (\mathbf{z}_{i,j} - \bar{\mathbf{z}}_i)(\mathbf{z}_{i,j} - \bar{\mathbf{z}}_i)^T, \quad (7)$$

sendo N_i o número de indivíduos de treinamento de cada classe i ($i = 1, 2$), o vetor $\mathbf{z}_{i,j}$ representa o indivíduo j da classe i , e os vetores $\bar{\mathbf{z}}_i$ e $\bar{\mathbf{z}}$ são a média do grupo i e a média global. A matriz \mathbf{Z} é calculada por meio da transformação da matriz de dados \mathbf{X} utilizando a matriz ϕ_{PCA} .

O principal objetivo do LDA é encontrar um hiperplano de projeção que maximize o critério de Fisher, definido pela razão do determinante de \mathbf{S}_b e \mathbf{S}_w , definido como:

$$\varphi_{lda} = \operatorname{argmax}_{\varphi} \frac{|\varphi^T \mathbf{S}_b \varphi|}{|\varphi^T \mathbf{S}_w \varphi|}, \quad (8)$$

tal que φ_{lda} é o vetor de transformação que projeta as amostras de um espaço p -dimensional para um espaço unidimensional (espaço do LDA).

O critério de Fisher é maximizado quando o hiperplano φ_{lda} é composto pelo autovetor principal de $\mathbf{S}_w^{-1} \mathbf{S}_b$ com autovalor correspondente não-nulo (FUKUNAGA, 2013; JOHNSON; WICHERN, 2006b).

Analogamente ao PCA, pode-se realizar a navegação no hiperplano φ_{lda} , gerando um mapa espacial das regiões mais discriminantes da face, saindo de um lado da dimensão de discriminância para o outro lado (DAVATZIKOS, 2004; SATO et al., 2008). Logo, é possível visualizar imagens que representem os fenômenos de dor e não dor capturados pelo LDA por meio da navegação neste hiperplano. Este método é descrito matematicamente como:

$$l = \bar{\mathbf{z}} + (j\sigma_i + \bar{\mathbf{z}}_i) \times \varphi_{lda}, \quad (9)$$

onde $j \in -2, -1, 0, 1, 2$, i é o grupo de amostragem, $\bar{\mathbf{z}}$ é a média global, e σ_i e $\bar{\mathbf{z}}_i$ são, respectivamente, o desvio padrão e a média de cada grupo plotado no espaço do LDA.

Porém, o poder do LDA pode decair caso p se aproxime do número de exemplos de treinamento. Observe que a matriz de covariância intra-classe S_w é uma função de no máximo $(N - 2)$ vetores linearmente independentes. Logo, o LDA é apropriado para casos em que o tamanho do conjunto de treinamento é maior que o número de características (SATO et al., 2008).

Para o caso das bases disponíveis de imagens faciais de recém-nascidos, o número de características acaba sendo maior que o número de exemplos de treinamento. Neste caso, S_w é uma matriz singular ou matematicamente instável para o cálculo de sua inversa. Logo, o LDA padrão não pode ser aplicado diretamente neste caso.

Para contornar este problema, foi utilizado o MLDA (*Maximum uncertainty LDA-base*) (THOMAZ, C. E. et al., 2004; THOMAZ; KITANI; GILLIES, 2006; THOMAZ, C. E. et al., 2007), que considera a questão de estabilizar a matriz intra-classe S_w com sua versão regularizada.

3.4 MÉTRICAS DE SIMILARIDADE

Dadas duas imagens I_{MOV} e I_{REF} , calcula-se a similaridade entre essas imagens por meio da somatória das distâncias de cada par de pixels correspondentes. Podemos também nos basear nos pontos fiduciais das imagens e calcular a diferença entre esses pontos, a partir da imagem móvel e da imagem de referência.

Para o cálculo dessas métricas, a área de interesse foi definida pela envoltória convexa formada pelos 68 pontos fiduciais.

Essas métricas, comumente utilizadas na literatura e descritas a seguir, são: MSE (*Mean Squared Error*), P_{CC} (*Pearson Correlation Coefficient*), NMI (*Normalized Mutual Information*) e ΣL^2 (Norma L^2) (MAINTZ; VIERGEVER, 1998).

3.4.1 Minimização do Erro Quadrático Médio

A minimização do erro quadrático médio é uma métrica de similaridade que calcula diretamente a diferença de intensidade entre os N elementos da imagem móvel I_{MOV} e os N elementos da imagem de referência I_{REF} , ou seja, esta métrica assume que as duas imagens tenham a mesma quantidade de pixels. O erro quadrático médio MSE pode ser calculado por:

$$MSE = \frac{1}{N} \sum_{i=1}^N (I_{MOV_i} - I_{REF_i})^2, \quad (10)$$

3.4.2 Coeficiente de Correlação de Pearson

A correlação é uma medida estatística que quantifica a dependência entre dois sinais distintos. Como uma imagem é um sinal composto pelas intensidades de cinza de seus N elementos, é possível medir a similaridade entre a imagem móvel I_{MOV} e a imagem de referência I_{REF} por meio da correlação entre as duas.

Esta métrica também pertence aos métodos estatísticos de registro de imagens, pois leva em consideração a intensidade média da imagem móvel ($\overline{I_{MOV}}$) e a intensidade média da imagem de referência ($\overline{I_{REF}}$).

Existem diversas formas de correlação, mas a correlação de Pearson é a medida mais conhecida para quantificar a dependência entre dois sinais (JOHNSON; WICHERN, 2006b), e pode ser calculada por:

$$P_{CC} = \frac{\sum_{i=1}^N (I_{MOV_i} - \overline{I_{MOV}})(I_{REF_i} - \overline{I_{REF}})}{\sqrt{\sum_{i=1}^N (I_{MOV_i} - \overline{I_{MOV}})^2 \sum_{i=1}^N (I_{REF_i} - \overline{I_{REF}})^2}}, \quad (11)$$

3.4.3 Informação Mútua Normalizada

A Informação Mútua (IM) mede quanta informação uma variável tem sobre a outra (neste caso as variáveis são as intensidades das imagens), em lugar de comparar as intensidades diretamente.

As entropias marginais das imagens I_{MOV} e I_{REF} podem ser descritas como:

$$H(I_{MOV}) = - \sum_{i=1}^N p(I_{MOV_i}) \log p(I_{MOV_i}), \quad (12)$$

$$H(I_{REF}) = - \sum_{i=1}^N p(I_{REF_i}) \log p(I_{REF_i}), \quad (13)$$

em que $p(I_{MOV_i})$ e $p(I_{REF_i})$ são as probabilidades dos pixels com intensidades I_{MOV_i} e I_{REF_i} ocorrerem nas imagens correspondentes. A informação mútua entre as imagens é dada por:

$$S_{IM}(I_{MOV}, I_{REF}) = H(I_{MOV}) + H(I_{REF}) - H(I_{MOV}, I_{REF}) \quad (14)$$

$$= - \sum_{i=1}^N p(I_{MOV_i}, I_{REF_i}) \log \frac{p(I_{MOV_i}, I_{REF_i})}{p(I_{MOV_i}) \cdot p(I_{REF_i})}, \quad (15)$$

em que $H(I_{MOV}, I_{REF})$ é definido como a Entropia Conjunta (EC) de $I_{MOV}(x)$ e $I_{REF}(x)$, descrita matematicamente como:

$$H(I_{MOV}, I_{REF}) = - \sum_{i=1}^N p(I_{MOV_i}, I_{REF_i}) \log p(I_{MOV_i}, I_{REF_i}), \quad (16)$$

a qual procura medir quão bem alinhadas essas imagens estão usando o conceito de entropia de *Shannon* (SHANNON, 2001; FABBRO, s.d.).

Pode-se observar pela Equação (15) que minimizar a entropia conjunta aumenta a informação mútua. Porém, a introdução das entropias marginais penaliza uma redução do montante de informação em cada imagem. Portanto, a IM é menos sensível ao alinhamento das imagens do que a entropia conjunta.

Embora a métrica IM seja menos sensível ao alinhamento das imagens do que a EC, esta ainda não apresenta a propriedade da invariância. Outra métrica que procura resolver as deficiências da IM é a Informação Mútua Normalizada (IMN) proposta por Studholme, Hille Hawkes (1999) e matematicamente descrita por:

$$NMI = \frac{H(I_{MOV}) + H(I_{REF})}{H(I_{MOV}, I_{REF})}, \quad (17)$$

3.4.4 Norma L^2

Ao contrário das métricas anteriores, que dependem de milhares de variáveis³, medir a similaridade por meio dos pontos fiduciais reduz o número de variáveis para manter apenas aquelas que são relevantes. Por conta disso, a qualidade do alinhamento entre duas imagens por meio dos pontos fiduciais acaba muitas vezes sendo superior, especialmente em imagens de face, em que o alinhamento não necessariamente deve ser feito pela intensidade do elemento (textura), e sim pela disposição entre esses (estrutura) (ZHU; RAMANAN, 2012).

Para uma imagem 2D, a correspondência entre os M pontos fiduciais da imagem móvel e da imagem de referência é calculada pela soma da distância entre eles. A transformação

³ considerando que uma imagem é composta por N pixels

ótima é aquela que minimiza esta soma, ou seja, a diferença mínima entre os pontos fiduciais da imagem.

A distância mais conhecida é a distância euclidiana (norma L^2), e a soma das distâncias euclidianas entre os pontos fiduciais das imagens pode ser expressa por:

$$\Sigma L^2 = \sum_{i=1}^M \sqrt{(x_{MOV_i} - x_{REF_i})^2 + (y_{MOV_i} - y_{REF_i})^2}, \quad (18)$$

3.5 MAPAS T

Por fim, para visualização das diferenças entre os resultados gerados para as bases COPE e UNIFESP, foram calculados mapas estatísticos denominadas aqui por mapas T, selecionando as imagens de cada extremo da navegação do MLDA. Para a construção destes mapas, primeiramente, é realizada uma diferença pixel a pixel dos pares de imagens, além de uma normalização para transformar todas essas diferenças em valores positivos, definida como:

$$m_{diff} = \sqrt{(I_1 - I_2)^2}, \quad (19)$$

em que I_1 é a imagem vetorizada do grupo 1 (dor) e I_2 a do grupo 2 (não-dor).

Computado este módulo da diferença, foi calculada a divisão pixel a pixel do vetor m_{diff} pelo desvio padrão global da base, tomando o cuidado de realizar este cálculo apenas nas pixels s , onde o desvio padrão não era zero, ou seja, $s < N$ e $\sigma_i(s) \neq 0$. Este processo pode ser descrito matematicamente como:

$$m_T(s) = m_{diff}(s) / \sigma(s) \quad (20)$$

sendo que neste processo m_T foi inicializado como um vetor de zeros, ou seja, ao final, as posições diferentes de s tinham o valor de zero.

4 EXPERIMENTOS E RESULTADOS

Este capítulo descreve os experimentos e resultados realizados nesta dissertação. Os resultados referentes às normalizações são apresentados nas 3 primeiras seções e os de extração multivariada de informação nas 3 finais.

4.1 ATLAS

Utilizando os métodos de normalização descritos no capítulo anterior, foram gerados os atlas de imagens de faces de recém-nascidos por meio do cálculo da média. Para a base de dados COPE, são apresentadas as imagens médias sem nenhum tipo de normalização aplicada às imagens na Figura 10a, normalizadas globalmente na Figura 10b, e normalizadas não rigidamente na Figura 10c. Por meio da mesma metodologia, obtemos os atlas para o banco de imagens desenvolvido pela UNIFESP, mostradas analogamente nas Figuras 11a, 11b e 11c.

Figura 10 – Atlas das imagens da COPE: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente

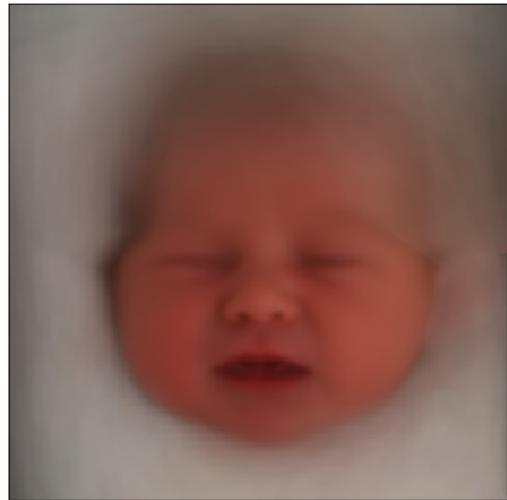
(a)



(b)



(c)



Fonte: Autor

Figura 11 – Atlas das imagens da UNIFESP: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente



Fonte: Autor

É possível visualizar a evolução alcançada por meio da metodologia de normalização proposta, dada a alta resolução dos detalhes faciais apresentados no atlas das imagens geradas. A normalização utilizando malhas triangulares nos proporciona uma imagem colorida com os detalhes de nariz, boca, olhos e contorno da face mais visíveis, a qual pode ser utilizada como referência em trabalhos que necessitem normalizar as bases apresentadas e que não possuam em mãos um atlas de alta resolução.

Com o intuito de se verificar e avaliar os resultados obtidos em relação aos atlas global e não-rígido, estes foram comparados qualitativamente por meio de imagens de desvio padrão (erro), além das métricas quantitativas de similaridade já destacadas no capítulo anterior.

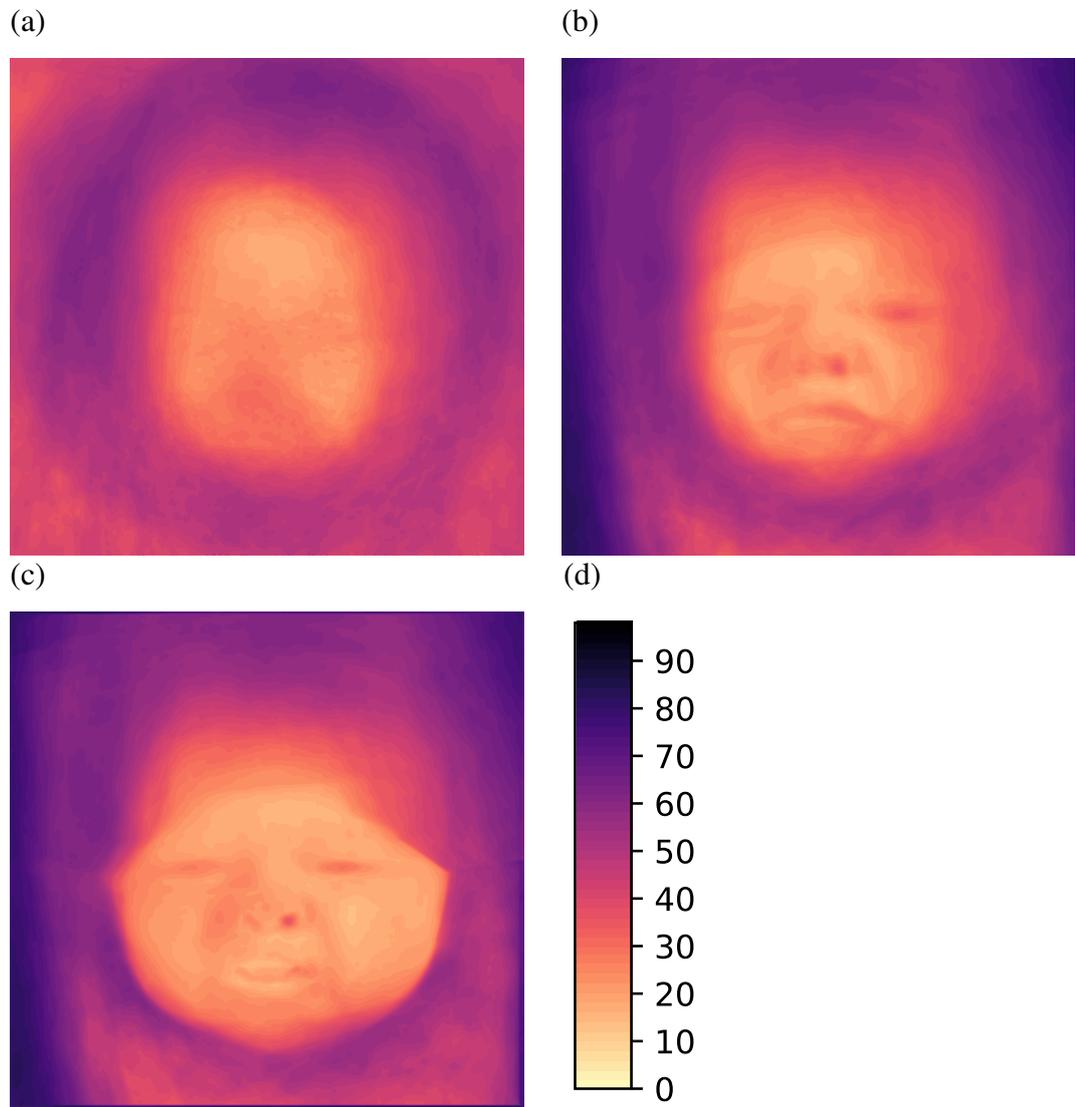
4.2 IMAGENS DE ERRO (QUALITATIVA)

Como visualizado anteriormente, para os atlas das imagens das bases COPE e UNIFESP, ocorre uma melhor definição dos detalhes faciais, além do contorno da face. Isso fica evidenciado nas Figuras 12 (COPE) e 13 (UNIFESP) em que temos valores muito maiores de desvio, principalmente em relação ao contorno, comparando-se a normalização global e a não-rígida, sendo que 0 é o valor mínimo de desvio padrão e 100 o valor máximo, apresentado nas Figuras 12d e 13d. As variações são ainda maiores quando comparamos os resultados obtidos pelas normalizações e o atlas das imagens originais.

Especificamente, analisando a Figura 12c, vemos valores diferentes de zero de desvio padrão em detalhes da face como olhos e narinas, e acredita-se que isso se dá por alguns fatores. Primeiramente, existem amostras dentro da base COPE que estão de olhos abertos e outras de olhos fechados, e isto é capturado pelo cálculo do desvio pois a normalização realizada irá apenas ajustar a amostra baseada em seus pontos fiduciais, ou seja, os pixels escuros da pupila serão posicionados para baixo, dado que na imagem de referência os olhos estejam fechados, gerando diferenças na região dos olhos quando compara-se amostras de olhos fechados e abertos. Para o caso das narinas, temos amostras em ângulos que não chegam a ilustra-las, pois a câmera se posicionou um pouco acima do bebê, e outras nas quais as narinas estão muito expostas, para o caso em que a câmera adquiriu uma imagem mais abaixo da face do bebê. Portanto, dadas essas diferenças, semelhante ao que ocorre com os olhos, isto fica evidenciado na imagem do desvio padrão, em que as narinas possuem valores de desvio maiores em relação aos outros detalhes da face.

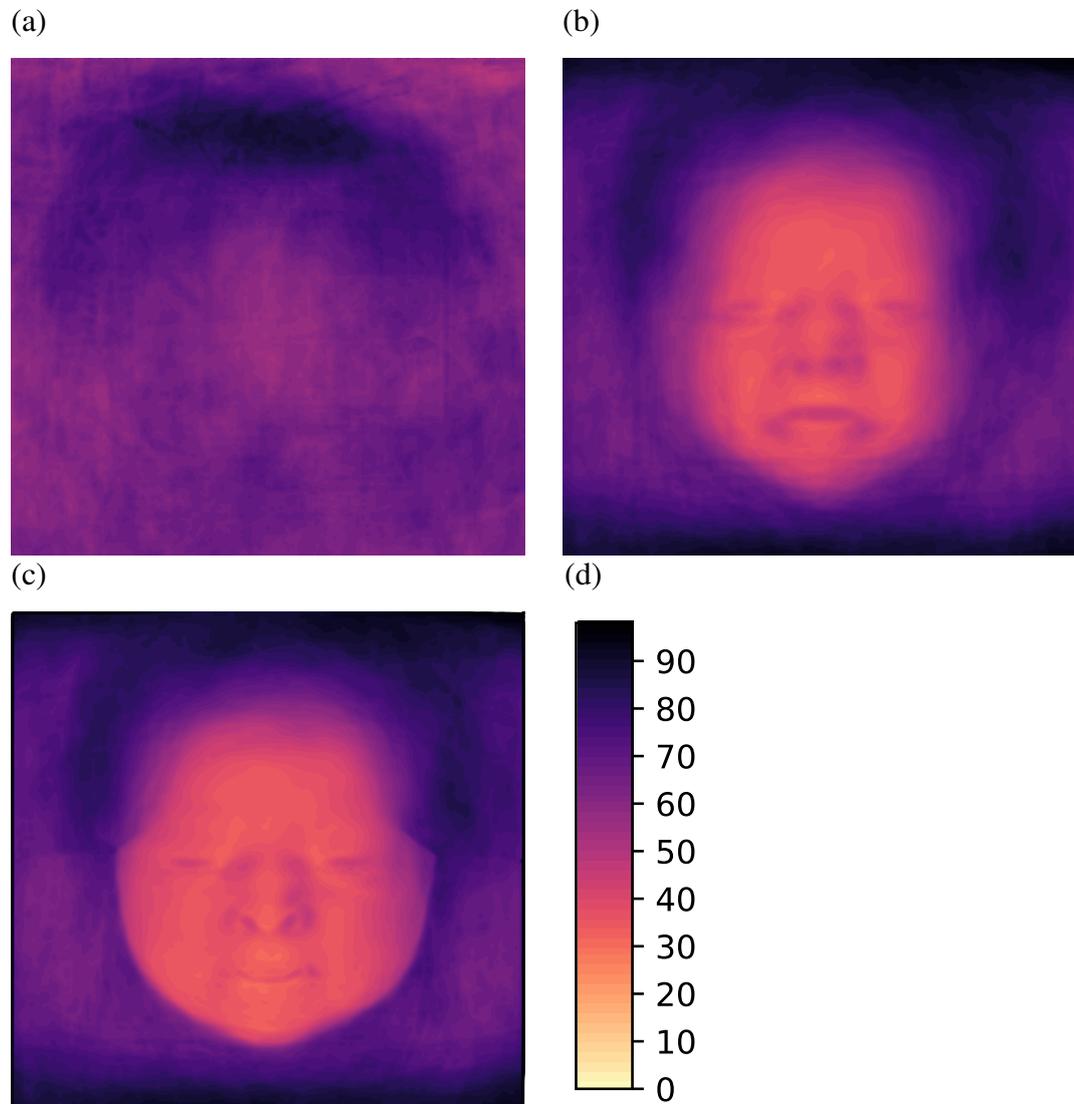
Adicionalmente, analisando a Figura 13c e tendo em mente que a base desenvolvida pela UNIFESP também apresenta as mesmas variações de angulação de câmera e estado dos olhos, podemos chegar às mesmas conclusões obtidas anteriormente para o desvio padrão do atlas das imagens normalizadas não rigidamente.

Figura 12 – Desvio padrão dos atlas das imagens da COPE: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente



Fonte: Autor

Figura 13 – Desvio padrão dos atlas das imagens da UNIFESP: (a) original, (b) normalizadas globalmente e (c) normalizadas não rigidamente



Fonte: Autor

4.3 MÉTRICAS DE SIMILARIDADE (QUANTITATIVA)

Além da análise qualitativa das imagens geradas pelo cálculo do desvio padrão, podemos fazer uso de métricas de similaridade que nos oferecem um resultado quantitativo para a avaliação dos atlas.

Foram utilizadas as métricas baseadas na diferença entre as intensidades dos pixels e também baseadas nos pontos fiduciais, como descrito no capítulo anterior, das imagens geradas através da normalização um-contra-todos, utilizando a imagem média global como referência.

Obtivemos os seguintes resultados:

Tabela 4 – Semelhança entre as imagens da base COPE

COPE — Métricas de similaridade			
	Original	Global	Não-rígida
MSE (↓)	639 ± 268	748 ± 329	129 ± 87
PCC (↑)	0.73 ± 0.08	0.79 ± 0.09	0.97 ± 0.02
NMI (↑)	0.44 ± 0.05	0.57 ± 0.05	0.68 ± 0.01
ΣL^2 (↓)	2668 ± 1099	519 ± 215	0.00 ± 0.00

Fonte: Autor

Tabela 5 – Semelhança entre as imagens da base da UNIFESP

UNIFESP — Métricas de similaridade			
	Original	Global	Não-rígida
MSE (↓)	2104 ± 856	1147 ± 597	379 ± 240
PCC (↑)	0.43 ± 0.21	0.82 ± 0.07	0.97 ± 0.02
NMI (↑)	0.33 ± 0.14	0.59 ± 0.05	0.70 ± 0.01
ΣL^2 (↓)	2776 ± 1175	464 ± 163	0.00 ± 0.00

Fonte: Autor

As Tabelas 4 e 5 nos fornecem mais uma comparação entre os atlas gerados para cada base de imagens. Analisando os valores obtidos para as métricas calculadas, a normalização não-rígida obteve, para ambos os conjuntos de dados, COPE e UNIFESP, os melhores resultados. Vale destacar o valor 0 alcançado com a normalização proposta utilizando a norma L^2 , como era de se esperar, pois o processo de triangulação força o reposicionamento dos pontos fiduciais de maneira que estes estejam exatamente na mesma posição quando comparada a imagem a ser normalizada com a de referência.

4.4 EXTRAÇÃO MULTIVARIADA DE INFORMAÇÃO

Deseja-se construir um hiperplano que consiga descrever e separar imagens faciais de dor e não-dor. No primeiro passo, foram extraídas as características mais relevantes de dor, e isso se deu por meio da aplicação do PCA. A seguir, foi utilizado o MLDA para separar os dados entre as duas classes. Para o PCA foram utilizadas o número de componentes não-nulas (FUKUNAGA, 2013), no caso da base COPE foram 116 e para a da UNIFESP foram 46.

4.4.1 SELEÇÃO DAS IMAGENS

Foram escolhidas apenas os *ground truths* para realizar o processo de normalização de todos-contratodos e construção dos hiperplanos, ou seja, somente as imagens rotuladas pelos especialistas com nota máxima de certeza como dor e não-dor na escala NFCS entraram nesta seleção. *Ground truth* é importante para classificação supervisionada das imagens, pois podemos garantir uma maior consistência do classificador construído, dado que estamos tratando de casos extremos, dor e não-dor, e retirando os casos de incerteza.

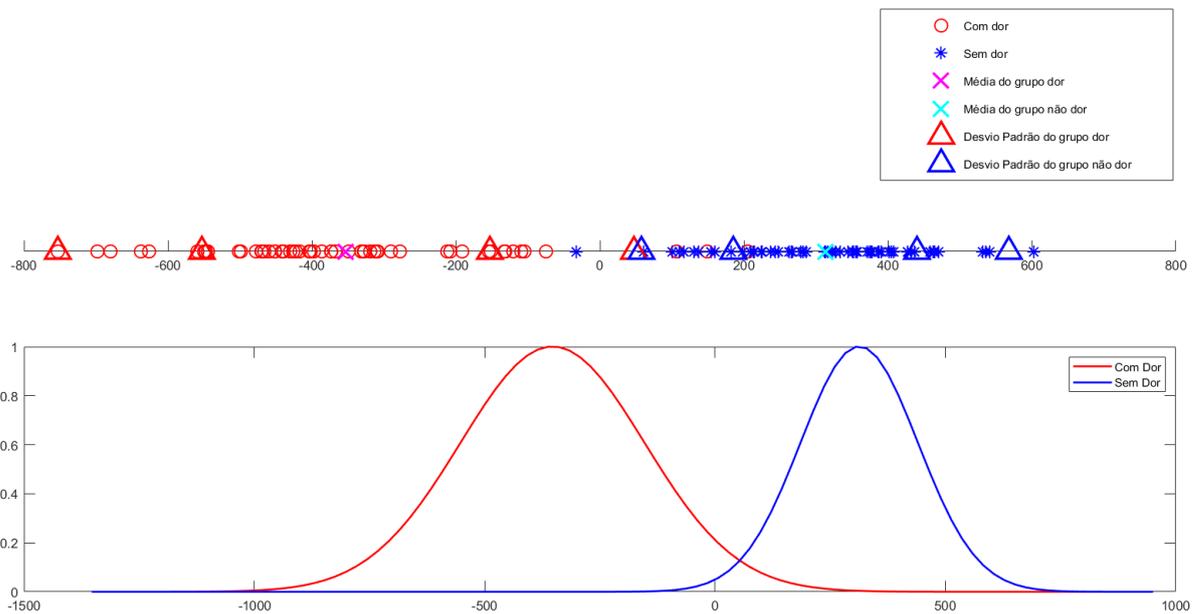
Foram também removidas da seleção final para construção do hiperplano as imagens que estavam prejudicadas de alguma maneira, seja por mal posicionamento da câmera no momento da captura, por artefatos posicionados sobre a face do recém-nascidos, entre outros. Vale considerar que o arcabouço construído aqui poderia sofrer prejudicialmente caso estas imagens entrassem para a normalização e posterior extração de informação, dado que não são tratadas, por exemplo, imagens com faces em ângulos muito diferentes do de perfil.

Por fim, para a base COPE, foram utilizadas 55 imagens de dor e 62 de não-dor, e para a UNIFESP foram utilizadas 21 imagens de dor e 26 de não-dor. Para o caso da base UNIFESP, foram considerados *ground truths* de dor apenas as imagens no instante da punção venosa, pois neste momento o recém-nascido está experienciando algum nível de dor.

4.4.2 CONSTRUÇÃO DO HIPERPLANO

Seguindo o processo descrito na seção 3.2, as imagens selecionadas foram normalizadas gerando um novo conjunto de dados. Foi então obtida para a base COPE a separação apresentada na Figura 14, em que os desvios padrões de cada classe foram multiplicados por i , onde $i \in \{-2, -1, 1, 2\}$. Observa-se que o algoritmo foi capaz de dividir no hiperplano discriminante as duas classes, dor e não-dor. Isso pode ser constatado também pelas gaussianas geradas por meio da separação.

Figura 14 – Projeção dos dados no eixo discriminante entre o grupo com dor (esquerda) e sem dor (direita) para a base COPE



Fonte: Autor

Para verificar as componentes geradas, foi calculada a navegação ao longo do vetor de componentes do PCA, de acordo com a Equação 5 descrita no capítulo anterior. A Figura 15 apresenta as 7 primeiras componentes de maior autovalor e suas respectivas navegações, sendo a imagem do meio a média global.

Figura 15 – Navegação no hiperplano das sete primeiras componentes principais, ordenadas da maior para menor variância, para a base COPE



Fonte: Autor

Sendo importante mencionar que o processo de normalização utilizando a metodologia de todos-contra-todos promove uma transformação menos agressiva dos dados, gerando regiões borradas principalmente no contorno da face.

Nesta imagem é possível encontrar alguns padrões de dor, como por exemplo, bocas abertas, olhos fechados e franzidos, entre outros. Também pode-se visualizar que a primeira componente provavelmente descreve artefatos globais das imagens, como iluminação e contraste, logo, pode-se realizar uma nova separação removendo-a, dado que esta componente provavelmente não descreve características relevantes para o fenômeno da dor.

Foi realizada também a navegação no hiperplano do MLDA, com o intuito de gerar imagens artificiais que representem uma atenuação ou mitigação da dor, gerando a Figura 16. Nesta, a imagem central se refere à média global e as restantes são geradas de acordo com a Equação 9, descrita no capítulo anterior, definindo os saltos da navegação.

Figura 16 – Navegação no hiperplano discriminante para a base COPE

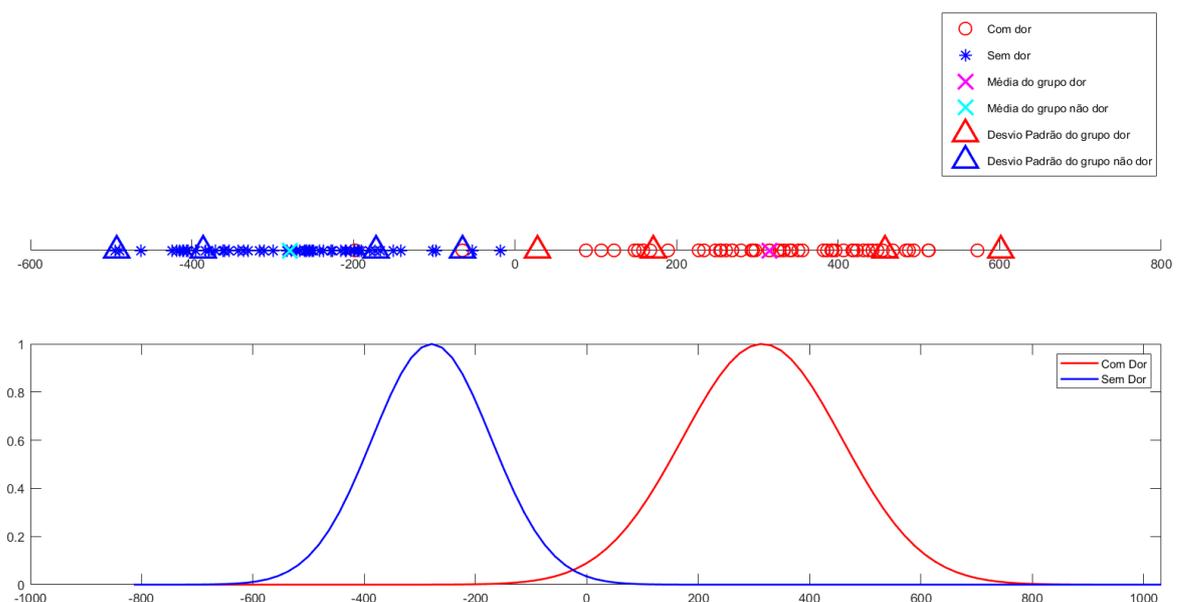


Fonte: Autor

Interessante notar que quanto mais caminha-se para esquerda, mais as características de dor começam a surgir, como por exemplo, testa franzida, olhos espremidos, aprofundamento do suco nasolabial e boca horizontalmente aberta. Em contrapartida, quando parte-se da imagem média global e caminha-se para a direita, características de não dor se tornam mais presentes, como fechamento da boca, abertura dos olhos, relaxamento da testa e etc.

Com o intuito de buscar uma separação ainda melhor dos dados, foi removida a primeira componente e gerado um novo hiperplano de separação, apresentado na Figura 17, junto com as respectivas gaussianas das duas classes.

Figura 17 – Projeção dos dados no eixo discriminante entre o grupo com dor e sem dor removendo a primeira componente principal para a base COPE



Fonte: Autor

Como visto na Figura 17, ocorrem menos sobreposições de imagens de classes distintas, sendo que, neste caso, indivíduos com dor foram posicionados no lado positivo do hiperplano, representado pela classe não-dor. Outro resultado interessante é que neste novo hiperplano, há uma distância maior entre os extremos das duas classes, sendo outro ponto positivo para esta separação, pois mostra que as classes estão melhor divididas.

Foi gerada também para este caso a navegação do hiperplano do MLDA, apresentada na Figura 18.

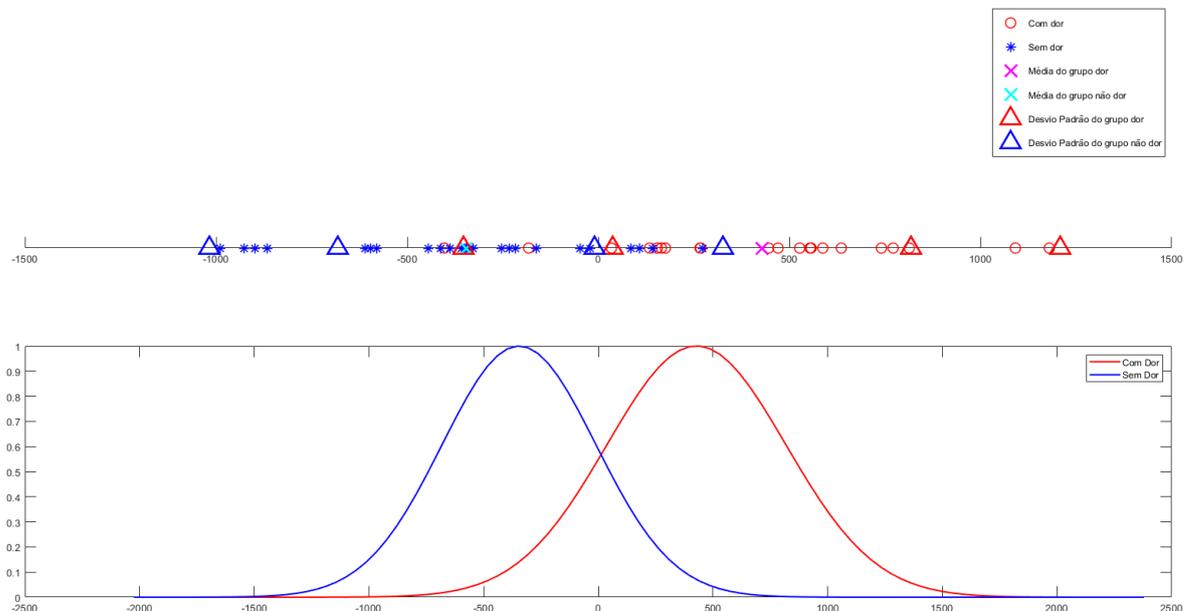
Figura 18 – Navegação no hiperplano discriminante removendo a primeira componente principal original do PCA para a base COPE



Fonte: Autor

Utilizando o mesmo procedimento descrito anteriormente, foi construído o hiperplano para separação das imagens da base UNIFESP entre dor e não-dor, apresentado na Figura 19.

Figura 19 – Projeção dos dados no eixo discriminante entre o grupo com dor (esquerda) e sem dor (direita) para a base UNIFESP



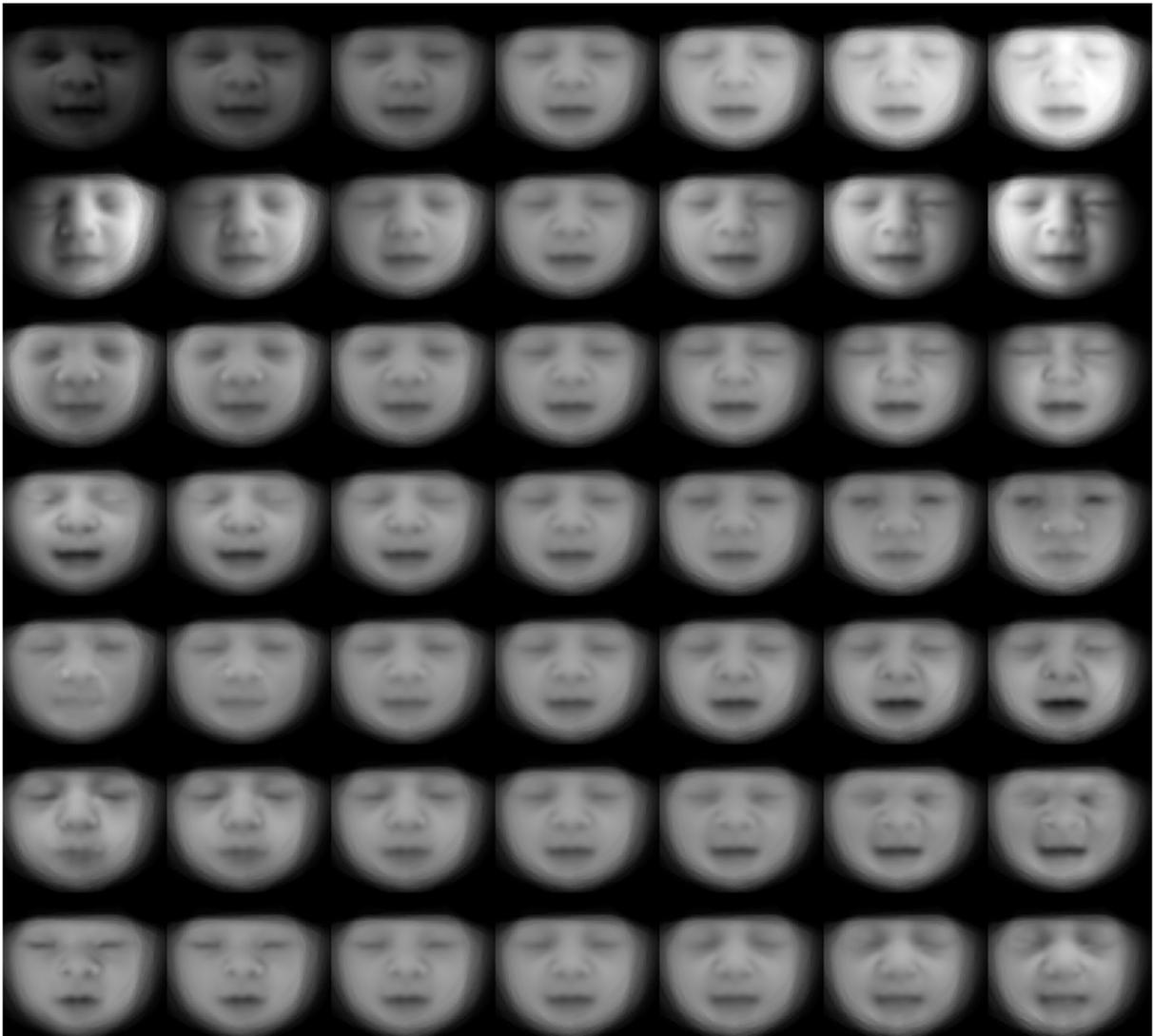
Fonte: Autor

Nota-se que, para esta base, a separação gerou um número maior de sobreposições entre os exemplos de cada classe; isso é ilustrado também nas gaussianas que, diferente das sepa-

rações da base COPE, apresentam uma sobreposição considerável, existindo uma região de dúvida entre dor e não-dor. Algumas suposições do porquê deste resultado seriam: o MLDA encontrou dificuldade em espalhar os indivíduos entre as regiões de separação linear, pois as imagens rotuladas como dor e não-dor podem, talvez, estar compartilhando características semelhantes, dificultando assim o posicionamento dos indivíduos no hiperplano; a quantidade de imagens não foi suficiente para o MLDA gerar um hiperplano assertivo. Note que para a base COPE o número de imagens foi maior que o dobro comparado com a UNIFESP.

Foram geradas também as navegações do PCA e do MLDA utilizando os vetores construídos no passo anterior, apresentadas nas Figuras 20 e 21, respectivamente.

Figura 20 – Navegação no hiperplano das sete primeiras componentes principais, ordenadas da maior para menor variância, para a base UNIFESP



Fonte: Autor

Figura 21 – Navegação no hiperplano discriminante para a base UNIFESP

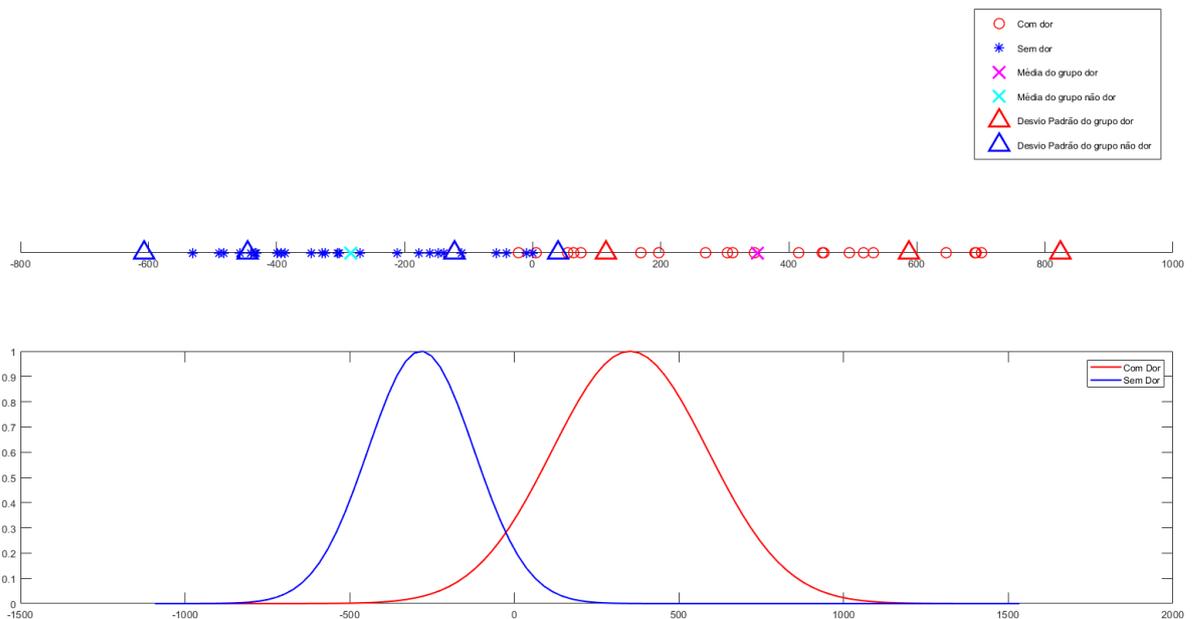


Fonte: Autor

Verifica-se também na análise das componentes principais que as duas primeiras componentes de maior autovalor não estão discriminando fenômenos relacionados à dor, e sim aos artefatos globais da imagem. A primeira parece descrever as variações de iluminação mais globais e a segunda variações de iluminação locais. Note que na segunda componente na imagem gerada no extremo esquerdo há uma iluminação no lado direito da face e na imagem gerada no extremo direito há uma iluminação no lado esquerdo da face, como se a componente tentasse descrever um objeto luminoso transladando ao redor da face do indivíduo.

Logo, foi aplicado novamente o PCA junto ao MLDA porém removendo essas duas componentes, além da navegação no hiperplano do MLDA, e os resultados são ilustrados nas Figuras 22 e 23.

Figura 22 – Projeção dos dados no eixo discriminante entre o grupo com dor e sem dor removendo a primeira componente principal para a base UNIFESP



Fonte: Autor

Figura 23 – Navegação no hiperplano discriminante removendo a primeira componente principal original do PCA para a base COPE



Fonte: Autor

Desta vez foi obtida uma melhor separação dos dados, ocorrendo poucos casos de sobreposição de indivíduos de classes distintas. Isso se reflete também nas gaussianas, onde a área de sobreposição é bem menor comparada à Figura 19.

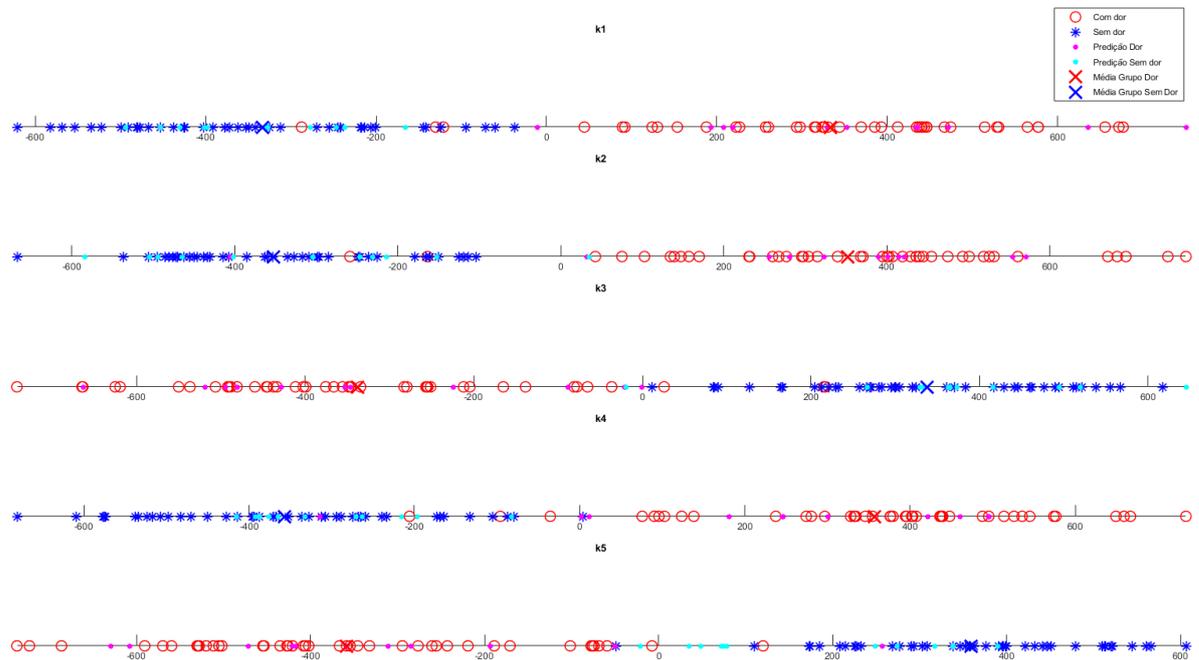
4.5 VALIDAÇÃO

Com o objetivo de avaliar os hiperplanos construídos, foram utilizados 2 métodos de validação cruzada, o *k-fold*, sendo seu resultado utilizado como limite inferior para a classificação, e o *leave-one-out*, utilizado como limite superior.

Seguindo o procedimento padrão de classificação, ambas as bases foram separadas em conjuntos de treinamento e teste. Todos os métodos de separação foram construídos utilizando a plataforma Matlab e a função *crossvalin* para separar aleatoriamente as imagens entre esses conjuntos. O *k-fold* foi implementado com $k = 5$, ou seja, foram realizadas 5 iterações sendo que 20% dos dados foram separados para teste e os 80% restantes para treinamento. Foi utilizada a distância euclidiana simples em relação ao indivíduo e às médias de cada grupo (dor e não-dor). Ao final da classificação, para cada conjunto de treinamento, foi realizado a média de todas as taxas de acerto de cada iteração. Considerando que os resultados de acurácia se referem à performance geral do modelo, ou seja, dentre todas as classificações (dor e não-dor), quantas ele classificou corretamente.

Primeiramente para a base COPE, utilizando *k-fold* e o hiperplano apresentado na Figura 16 com todas as componentes do PCA, foi obtida uma acurácia média de 83,64% e as separações descritas na Figura 24:

Figura 24 – K-fold com $k = 5$ utilizando o primeiro hiperplano discriminante construído para a base COPE

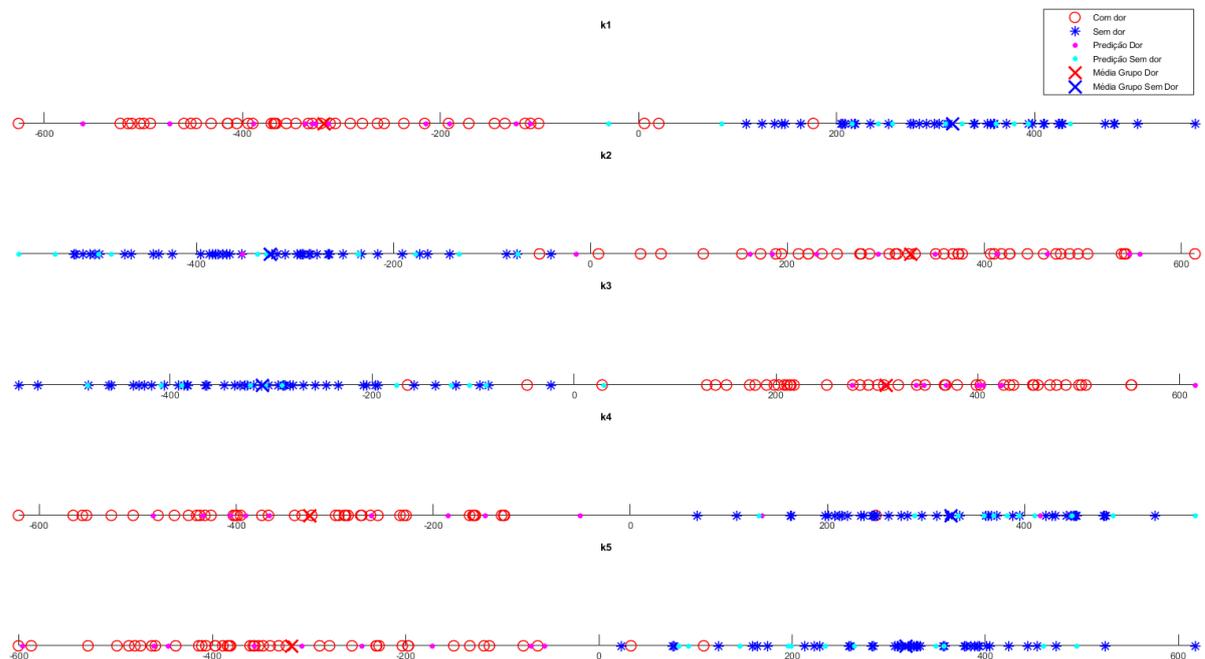


Fonte: Autor

Já para o *leave-one-out* foi alcançada a acurácia média de 95,73%, observando que houve um aumento considerável de acerto, porém devemos considerar que este é o valor máximo de classificação que o arcabouço construído pode alcançar, utilizando o hiperplano em questão.

Utilizando o segundo hiperplano construído utilizando a base COPE e removendo a primeira componente do PCA (Figura 18), foi aplicado novamente o *k-fold* com $k = 5$, e obtida uma acurácia média de 87,27%, maior do que a anterior que utilizou o primeiro hiperplano. As separações obtidas neste processo são apresentadas na Figura 25:

Figura 25 – K-fold com $k = 5$ utilizando o segundo hiperplano discriminante construído para a base COPE

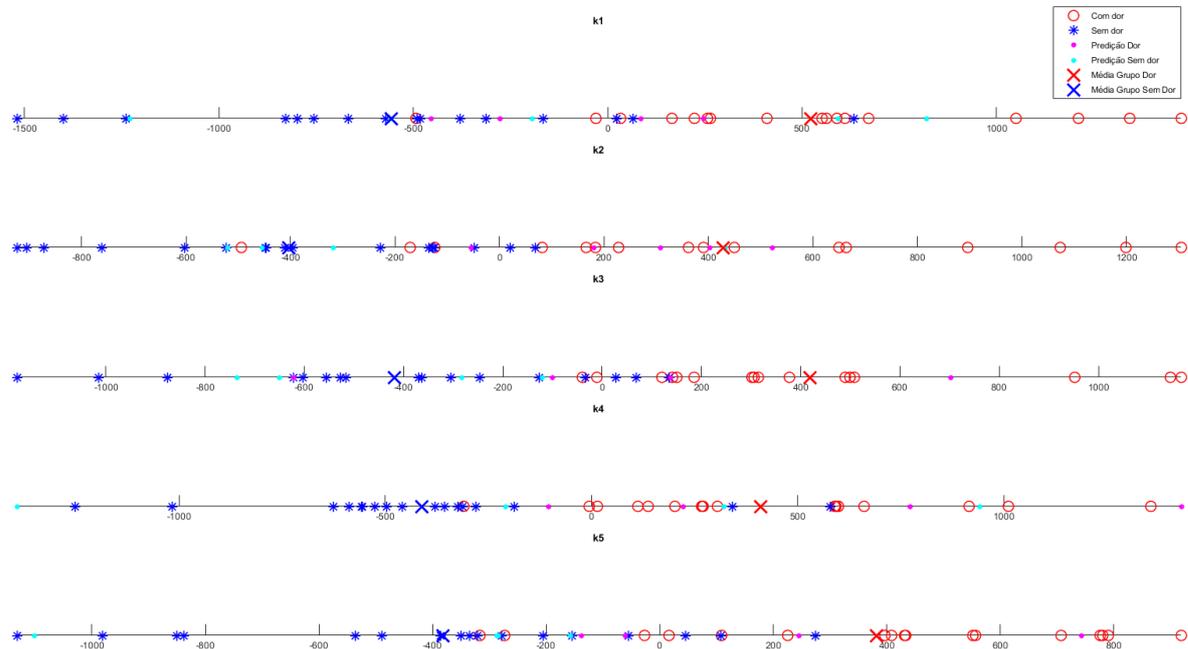


Fonte: Autor

Finalmente para o *leave-one-out* foi alcançada a acurácia média de 95,73%, valor exatamente igual ao obtido utilizando o primeiro hiperplano. Portanto, para a base COPE, foi obtida uma acurácia relativamente alta, comprovando que o hiperplano construído, principalmente o segundo, pode ser utilizado para classificar imagens faciais de bebês recém-nascidos entre dor e não-dor.

Em seguida, utilizando a base UNIFESP para o processo de classificação, foi aplicado o *k-fold* com $k = 5$ e o hiperplano apresentado na Figura 21, foi obtida a acurácia média de 64,17%, bem menor em comparação à acurácia média obtida utilizando a base COPE, dado que a separação realizada pelo MLDA não foi tão assertiva. As separações obtidas em cada iteração do *k-fold* são apresentadas na Figura 26.

Figura 26 – K-fold com $k = 5$ utilizando o primeiro hiperplano discriminante construído para a base UNIFESP.

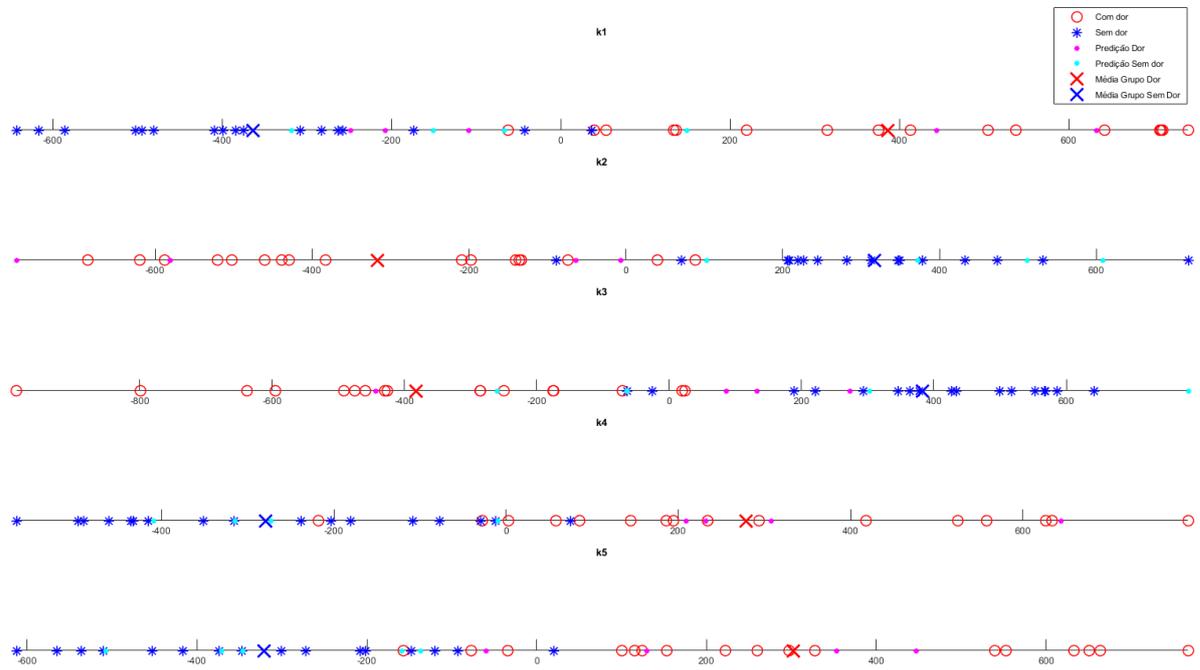


Fonte: Autor

Já aplicando o *leave-one-out* foi alcançada a acurácia média de 65,96%, ou seja, para a base UNIFESP, utilizando o primeiro hiperplano gerado, chegou-se à um limite inferior de 64,17% e um superior de 65,96%.

Finalmente, utilizando o segundo hiperplano, onde foram removidas as duas componentes principais de maior autovalor, chegou-se à uma acurácia média de 68,89% utilizando o *k-fold* com $k = 5$. As separações obtidas em cada iteração do *k-fold* são apresentadas na Figura 27.

Figura 27 – K-fold com $k = 5$ utilizando o segundo hiperplano discriminante construído para a base UNIFESP



Fonte: Autor

Neste caso, foi obtida uma acurácia média de 76,6% para o *leave-one-out*. Esta foi a maior taxa alcançada entre os métodos, embora este seja um valor relativamente baixo. Supõe-se novamente as mesmas questões discutidas nos resultados obtidos na construção do primeiro hiperplano da base UNIFESP.

As Tabelas 6 e 7 resumem todos os resultados obtidos para as bases COPE e UNIFESP, respectivamente, com as informações de acurácia, sensibilidade, especificidade, *precision* e *recall*, sendo que para o k-fold foi calculada a média e o desvio padrão dessas métricas:

Tabela 6 – Resultados das classificações para a base COPE

COPE — K-Fold		
	PCA + MLDA	PCA (1 componente a menos) + MLDA
Acurácia	0,8364 ± 0,0689	0,8727 ± 0,0874
Sensitividade	0,7758 ± 0,1073	0,8292 ± 0,1065
Especificidade	0,9306 ± 0,0636	0,9500 ± 0,0685
Precision	0,9457 ± 0,0512	0,9652 ± 0,0478
Recall	0,7758 ± 0,1073	0,8292 ± 0,1065
COPE — Leave-One-Out		
	PCA + MLDA	PCA (1 componente a menos) + MLDA
Acurácia	0,9573	0,9573
Sensitividade	0,9839	0,9839
Especificidade	0,9273	0,9273
Precision	0,9385	0,9385
Recall	0,9839	0,9839

Fonte: Autor

Tabela 7 – Resultados das classificações para a base UNIFESP

UNIFESP — K-Fold		
	PCA + MLDA	PCA (2 componentes a menos) + MLDA
Acurácia	0,6417 ± 0,1782	0,6889 ± 0,0712
Sensitividade	0,6267 ± 0,2477	0,7414 ± 0,1991
Especificidade	0,6667 ± 0,2041	0,6633 ± 0,2329
Precision	0,7533 ± 0,1850	0,7857 ± 0,1577
Recall	0,6267 ± 0,2477	0,7414 ± 0,1991
UNIFESP — Leave-One-Out		
	PCA + MLDA	PCA (2 componentes a menos) + MLDA
Acurácia	0.6596	0.7660
Sensitividade	0.7308	0.8462
Especificidade	0.5714	0.6667
Precision	0.6786	0.7586
Recall	0.7308	0.8462

Fonte: Autor

Para a base COPE, temos resultados consideráveis comparados com os anteriores que utilizaram a mesma base, destacados no capítulo 2 de revisão da literatura, sendo que devemos levar em consideração que o limite inferior alcança acurácias bem próximas aos resultados de outros trabalhos (BRAHNAM et al., 2006, 2007; GHOLAMI; HADDAD; TANNENBAUM, 2010; NANNI; BRAHNAM; LUMINI, 2010; MANSOR; REJAB, 2013b; BRAHNAM; NANNI; SEXTON, R., 2007; BRAHNAM; NANNI; SEXTON, R. S., 2008; MANSOR; REJAB, 2013a), e o limite superior de aproximadamente 96% alcança porcentagens de acurácia

maiores que os apresentados em trabalhos anteriores, com a ressalva de que foi realizada uma separação de imagens anterior à normalização e classificação.

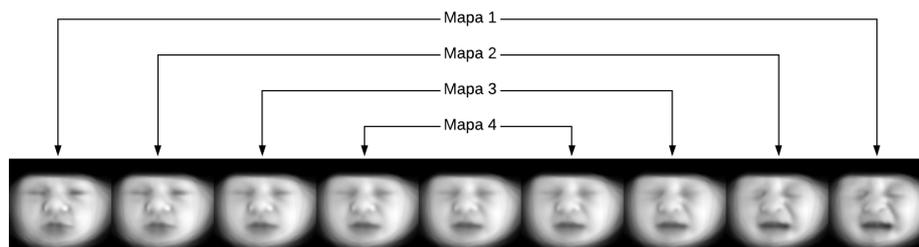
Já para a base UNIFESP, os resultados não são tão positivos quanto os da base COPE, porém, considerando todos os aspectos já citados anteriormente em relação às dificuldades em se normalizar esta base e o número reduzido de imagens, gerando um hiperplano não tão assertivo, temos valores de acurácia consideráveis, mas que podem ser melhorados por meio de uma nova seleção de imagens ou utilizando-se diferentes métricas de avaliação dos hiperplanos gerados nos métodos de validação.

4.6 MAPAS T DE DOR

Foi realizada a comparação entre as imagens artificiais geradas na navegação do hiperplano do MLDA através da construção dos Mapas T, descrita no capítulo anterior, para verificar quais regiões faciais se destacam nos momentos em que o recém-nascido está sofrendo experiências dolorosas.

Por meio deste processo, foi utilizada a seguinte seleção de pares apresentada na Figura 28, sendo que 0 significa a diferença normalizada mínima entre as imagens e 100 a máxima.

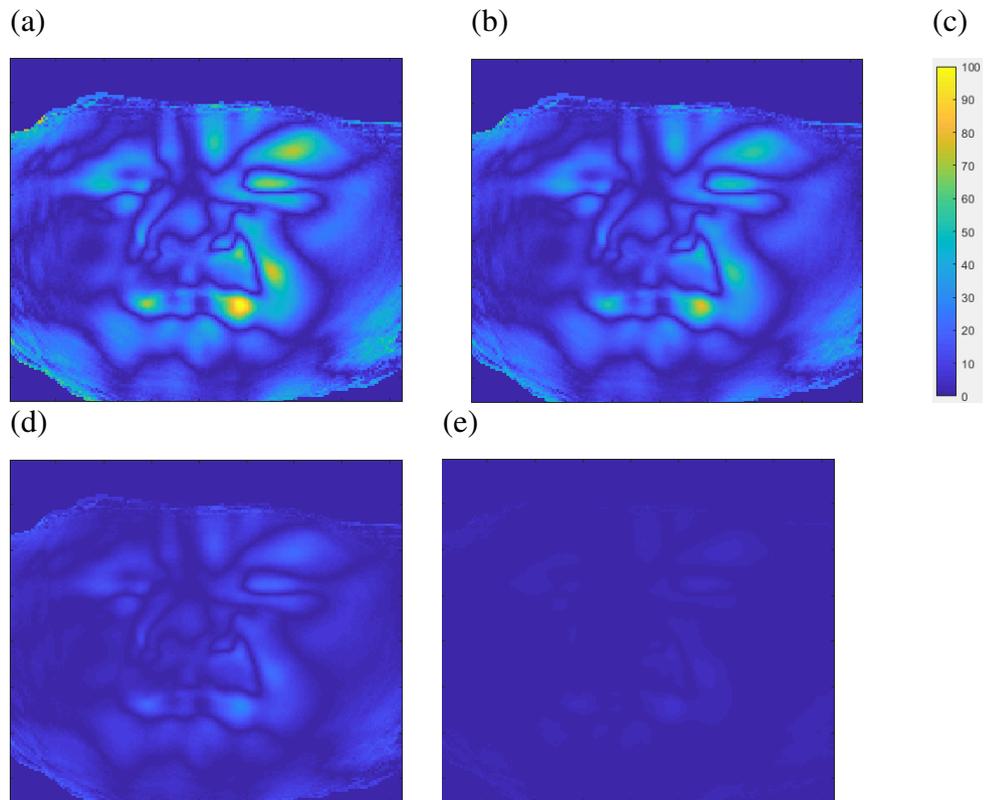
Figura 28 – Seleção de pares para construção dos mapas



Fonte: Autor

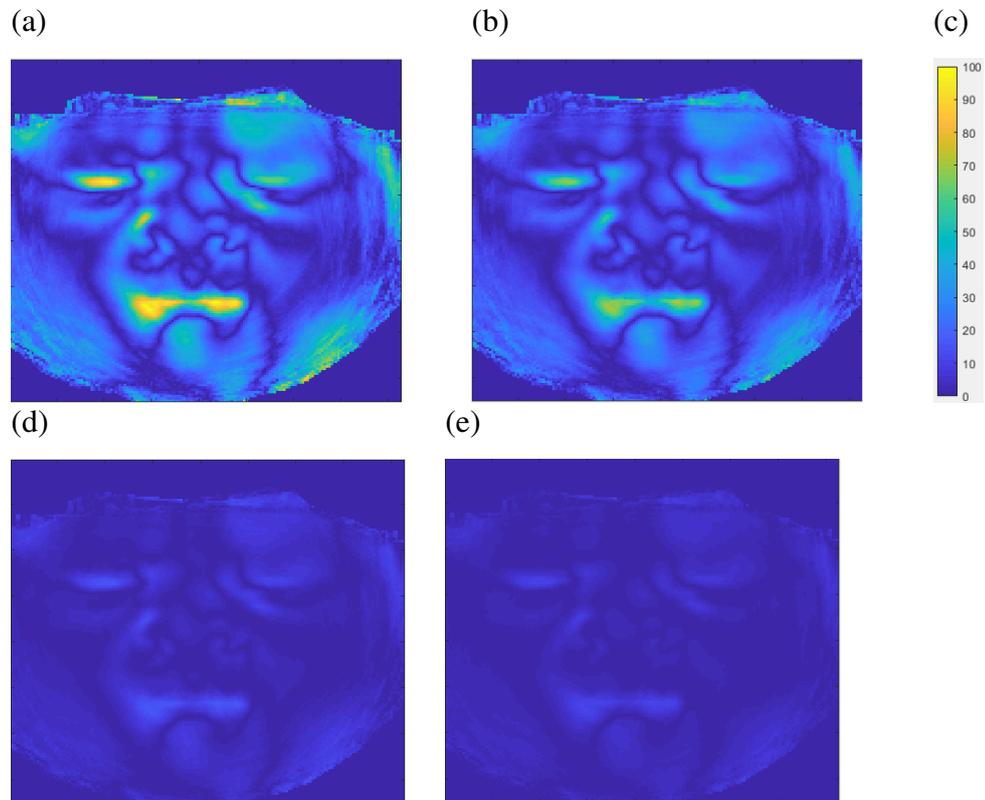
Com as imagens devidamente selecionadas, foram gerados os mapas para as bases COPE e UNIFESP, apresentadas nas Figuras 29 e 30, respectivamente.

Figura 29 – Mapas T das imagens geradas na navegação do hiperplano do MLDA para a base COPE: (a) mapa 1, (b) mapa 2, (c) mapa 3 e (d) mapa 4



Fonte: Autor

Figura 30 – Mapas T das imagens geradas na navegação do hiperplano do MLDA para a base UNIFESP (a) mapa 1, (b) mapa 2, (c) mapa 3 e (d) mapa 4

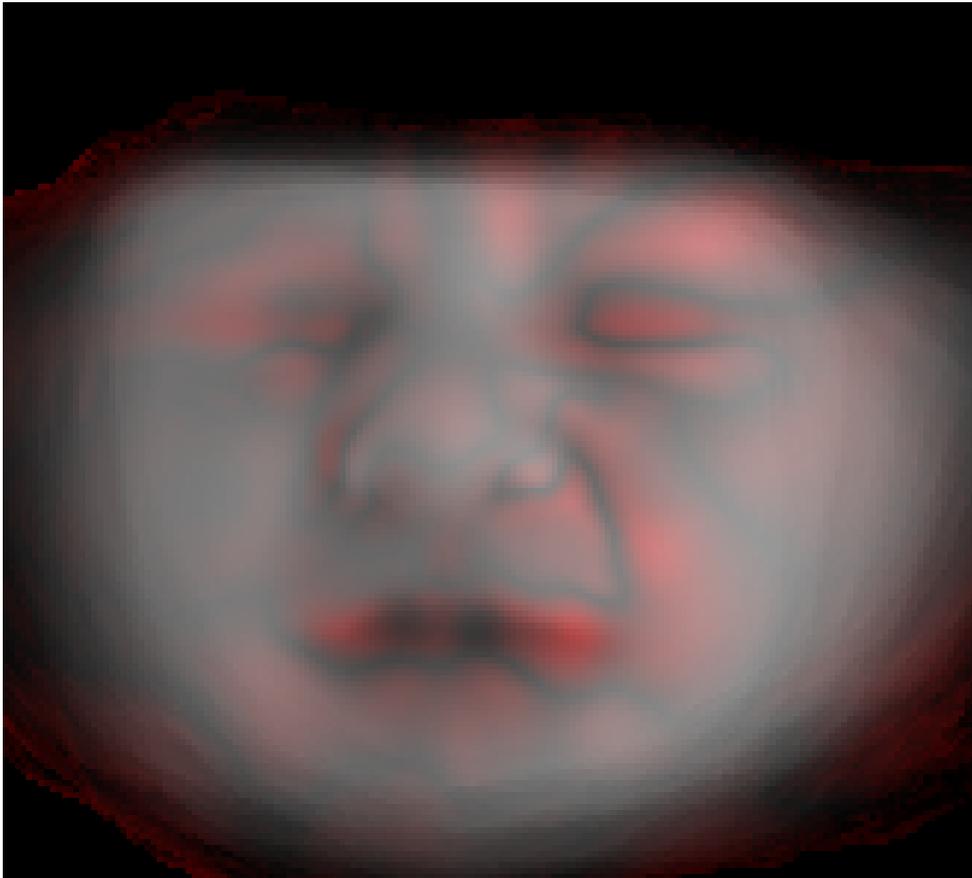


Fonte: Autor

Com o objetivo de identificar as regiões da face mais destacadas no momento de uma experiência dolorosa, primeiramente foi aplicado nos mapas um filtro da cor vermelha, mais precisamente uma multiplicação dos pixels da imagem original dos mapas por uma imagem de mesmo tamanho e cor sólida vermelha, sendo o resultado final da transformação, um mapa em que as regiões de maior destaque receberam tons de vermelho mais vivos, e as de menos destaque tons mais claros. Após isto, foi aplicada uma sobreposição dos mapas das Figuras 29a e 30a, já transformados em tons de vermelho, sobre os atlas médios de ambas as bases.

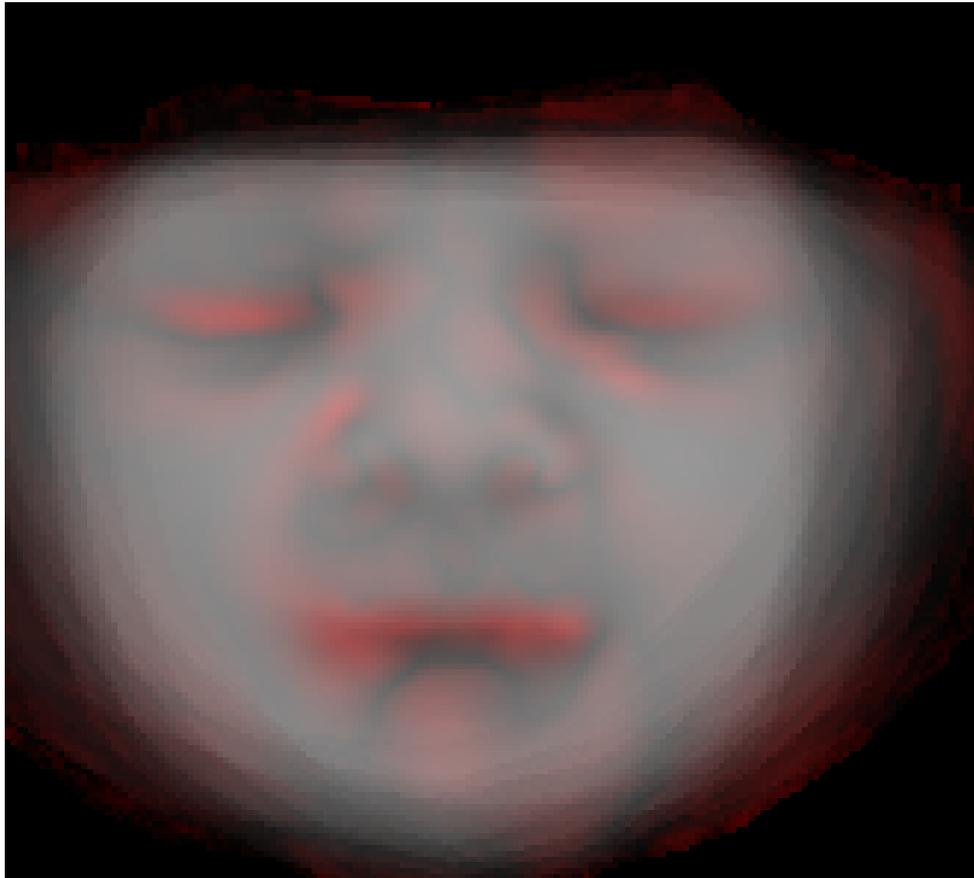
Por meio deste processo, foram geradas as imagens das Figuras 31 e 32 para as bases COPE e UNIFESP, respectivamente.

Figura 31 – Regiões de destaque para a base COPE



Fonte: Autor

Figura 32 – Regiões de destaque para a base UNIFESP



Fonte: Autor

Analisando as Figuras 31 e 32, vemos que o entorno da boca, quase toda a região orbitária, uma porção da região jugal (da bochecha), além da elevação das sobrancelhas e enrijecimento da glabella são partes da face que recebem destaque em detrimento das outras em momentos que um recém nascido está experienciando dor, segundo os mapas T gerados por meio das navegações do MLDA.

Desconsiderando as áreas externas à face, os pontos de destaque especificados anteriormente servem como informação estatística discriminante para profissionais da saúde, pois estes poderiam prestar uma maior atenção a estas regiões e provavelmente serem mais assertivos quanto a avaliação de dor em momentos de urgência. Estas imagens podem ser úteis também para a criação de algoritmos que fossem projetados de maneira a avaliar dor segundo essas regiões de destaque, ou seja, toda o processamento e análise recairia nestas regiões, ou então pesos maiores poderiam ser atribuídos a essas áreas. Há a possibilidade de se criar uma nova escala de dor, ou enriquecer as existentes, utilizando estas regiões como pontos de destaque.

5 CONCLUSÃO

Esta dissertação descreveu um arcabouço computacional de reconhecimento de padrões faciais utilizando passos bem definidos de normalização e extração multivariada de informação. Para a primeira etapa, foi construído um processo de alinhamento global e não-rígido utilizando malhas triangulares, gerando atlas realistas de face não-enviesados com um alto grau de detalhamento. Este processo apresentou como saída novas imagens normalizadas, removendo artefatos que poderiam atrapalhar e confundir o classificador no processo de identificação de dor, além de discrepâncias envolvendo angulação e posicionamento, ajustando as faces para uma posição frontal. Neste processo foram eliminados efeitos ou artefatos da imagem que não eram relevantes e que retratavam particularidades indesejadas inerentes a todo processo imperfeito de aquisição de dados, garantindo que estes não fossem propagados erroneamente como variações discriminantes.

Subsequentemente, foram avaliados hiperplanos estatísticos gerados através da combinação PCA+MLDA, aplicada às bases de imagens faciais de recém-nascidos investigadas (COPE e UNIFESP), mostrando que podemos separar estas bases entre imagens de dor e não-dor utilizando estes métodos lineares. Os resultados obtidos através do arcabouço construído nesta dissertação podem ser comparados com os apresentados na literatura atual, dado que para a base COPE a taxa de acurácia utilizando *leave-one-out* foi de mais de 95%, ultrapassando os valores obtidos até então utilizando esta mesma base de imagens. Porém, deve-se ter em mente que o objetivo deste trabalho não é apenas apresentar boas taxas de acerto para a classificação de dor, mas principalmente permitir a interpretação e o reconhecimento quantitativo de padrões faciais para avaliação de dor para profissionais de saúde. Por este motivo, foram gerados os mapas T, pois estes descrevem quantitativamente as regiões mais discriminantes quando está se discutindo avaliação de dor em imagens faciais, como por exemplo, regiões próximas aos olhos e boca discriminam muito mais dor que o restante da face. Portanto, estes mapas podem auxiliar os profissionais a realizarem uma avaliação mais específica e assertiva sobre a atual situação do recém-nascido.

Entretanto, deve-se destacar as seguintes limitações com relação ao arcabouço proposto. Foi utilizada uma seleção específica anterior à normalização e classificação, escolhendo apenas as imagens com ângulos próximos ao de perfil, sendo que pode-se realizar uma nova seleção, adicionando ou removendo imagens com diferentes angulações. Os melhores resultados foram alcançados removendo as componentes do PCA que aparentemente descreviam mudanças glo-

bais da imagem, como luz, sombra e etc, e essa seleção de componentes foi feita por meio da análise visual dos resultados gerados na caminhada do PCA. É possível, portanto, realizar uma escolha mais assertiva utilizando algoritmos específicos para seleção das componentes que mais discriminam dor. Neste trabalho foi utilizada a distância euclidiana simples em relação ao indivíduo e às médias de cada grupo (dor e não-dor) para determinar a classe final no momento da predição, porém outros classificadores podem ser utilizados neste caso, e talvez porcentagens de acurácia maiores possam ser alcançadas. Para os casos de bases com poucas imagens, como é o caso da UNIFESP, pode-se utilizar imagens artificiais gerados no processo de navegação (PCA+MLDA) para enriquecer a base e talvez obter melhores porcentagens de acurácia. E por fim, ambas as bases foram construídas utilizando bebês a termo, ou seja, ainda existe o desafio de avaliar imagens de recém-nascidos que se encontram em situações adversas, como por exemplo, bebês que necessitam de aparelhos extras no processo de recuperação dentro de uma UTI neonatal.

A partir desta dissertação, algumas possibilidades, que vão além do que foi apresentado aqui, podem ser executadas como trabalhos futuros, tais como: Utilizar o arcabouço construído como *backend* de um sistema que monitora bebês recém-nascidos e notifica através de mensagens momentos de possíveis experiências de dor por parte dos indivíduos; Testar novos algoritmos de classificação utilizando o mesmo método de normalização, como, por exemplo, KNN, árvores de decisão, ou redes neurais.

Por fim, espera-se que com este trabalho a classificação de dor em recém-nascidos utilizando imagens faciais tenha dado mais um passo frente à assertividade e melhor compreensão clínica do fenômeno dor.

REFERÊNCIAS

- ALJABAR, Paul et al. Assessment of brain growth in early childhood using deformation-based morphometry. **Neuroimage**, Elsevier, v. 39, n. 1, p. 348–358, 2008.
- AMBUEL, Bruce et al. Assessing distress in pediatric intensive care environments: the COMFORT scale. **Journal of pediatric psychology**, Oxford University Press, v. 17, n. 1, p. 95–109, 1992.
- ANAND, Kanwaljeet JS; CRAIG, Kenneth D. New perspectives on the definition of pain. **Pain-Journal of the International Association for the Study of Pain**, [Amsterdam]: Elsevier/North-Holland, 1975-, v. 67, n. 1, p. 3–6, 1996.
- ANAND, Kanwaljeet JS; STEVENS, Bonnie J; MCGRATH, Patrick J (Ed.). **Pain in neonates and infants**. 3. ed. [S.l.]: Elsevier Health Sciences, 2007. cap. 6, p. 68–70.
- ARIAS, Maria Carmenza Cuenca; GUINSBURG, Ruth. Differences between uni-and multidimensional scales for assessing pain in term newborn infants at the bedside. **Clinics**, SciELO Brasil, v. 67, n. 10, p. 1165–1170, 2012.
- BALDA, Rita de Cássia Xavier et al. The recognition of facial expression of pain in full-term newborns by parents and health professionals. **Archives of pediatrics & adolescent medicine**, American Medical Association, v. 154, n. 10, p. 1009–1016, 2000.
- BARBER, C. Bradford; DOBKIN, David P.; HUHDANPAA, Hannu. The Quickhull algorithm for convex hulls. **ACM TRANSACTIONS ON MATHEMATICAL SOFTWARE**, v. 22, n. 4, p. 469–483, 1996.
- BELLIENI, Carlo V et al. Development and validation of the ABC pain scale for healthy full-term babies. **Acta Paediatrica**, Wiley Online Library, v. 94, n. 10, p. 1432–1436, 2005.
- BLAUER, Teresa; GERSTMANN, Dale. A simultaneous comparison of three neonatal pain scales during common NICU procedures. **The Clinical journal of pain**, LWW, v. 14, n. 1, p. 39–47, 1998.
- BRAHNAM, Sheryl; NANNI, Loris; SEXTON, Randall. Introduction to neonatal facial pain detection using common and advanced face classification techniques. In: BRAHNAM, Sheryl; NANNI, Loris; SEXTON, Randall (Ed.). **Advanced Computational Intelligence Paradigms in Healthcare**. [S.l.]: Springer, 2007. p. 225–253.
- BRAHNAM, Sheryl; NANNI, Loris; SEXTON, Randall S. Neonatal Facial Pain Detection Using NNSOA and LSVM. In: ARABNIA, Hamid R. (Ed.). **Proceedings of the 2008 International Conference on Image Processing, Computer Vision, & Pattern Recognition, IPCV**. [S.l.: s.n.], 2008. p. 352–357.
- BRAHNAM, Sheryl et al. Machine assessment of neonatal facial expressions of acute pain. **Decision Support Systems**, Elsevier, v. 43, n. 4, p. 1242–1254, 2007.
- BRAHNAM, Sheryl et al. Machine recognition and representation of neonatal facial displays of acute pain. **Artificial intelligence in medicine**, Elsevier, v. 36, n. 3, p. 211–222, 2006.
- BRAHNAM, Sheryl et al. SVM classification of neonatal facial images of pain. In: BRAHNAM, Sheryl et al. (Ed.). **Fuzzy Logic and Applications**. [S.l.: s.n.], 2005. Springer, p. 121–128.
- BUCHHOLZ, Mark et al. Pain scores in infants: a modified infant pain scale versus visual analogue. **Journal of pain and symptom management**, Elsevier, v. 15, n. 2, p. 117–124, 1998.

- BULAT, A.; TZIMIROPOULOS, G. How Far are We from Solving the 2D & 3D Face Alignment Problem? (and a Dataset of 230,000 3D Facial Landmarks). In: 2017 IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV). Los Alamitos, CA, USA: IEEE Computer Society, out. 2017.
- CHERMONT, Aurimery Gomes et al. O que os pediatras conhecem sobre avaliação e tratamento da dor no recém-nascido? **Jornal de Pediatria**, SciELO Brasil, 2003.
- COOTES, Timothy F.; EDWARDS, Gareth J.; TAYLOR, Christopher J. Active appearance models. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 23, n. 6, p. 681–685, 2001.
- CRAIG, Kenneth D et al. Pain in the preterm neonate: behavioural and physiological indices. **Pain**, Elsevier, v. 52, n. 3, p. 287–299, 1993.
- DAVATZIKOS, Christos. Why voxel-based morphometric analysis should be used with great caution when characterizing group differences. **Neuroimage**, Elsevier, v. 23, n. 1, p. 17–20, 2004.
- DEBILLON, T et al. Development and initial validation of the EDIN scale, a new tool for assessing prolonged pain in preterm infants. **Archives of Disease in Childhood-Fetal and Neonatal Edition**, BMJ Publishing Group, v. 85, n. 1, f36–f41, 2001.
- FABBRO, Davi Araujo Dal. **Normalização espacial não rígida de imagens assistida por metamodelos**. Orientador: Carlos Eduardo Thomaz, 2018. 59 f. Dissertação (Mestrado em Engenharia Elétrica) – Centro Universitário FEI, São Bernardo do Campo, 2018.
- FUKUNAGA, Keinosuke. **Introduction to statistical pattern recognition**. [S.l.]: Elsevier, 2013.
- GHOLAMI, Behnood; HADDAD, Wassim M; TANNENBAUM, Allen R. Agitation and pain assessment using digital imaging. In: ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, 2009. EMBC 2009. ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE. Minneapolis, MN, USA: IEEE, 2009. p. 2176–2179.
- GHOLAMI, Behnood; HADDAD, Wassim M; TANNENBAUM, Allen R. Relevance vector machine learning for neonate pain intensity assessment using digital imaging. **IEEE Transactions on biomedical engineering**, IEEE, v. 57, n. 6, p. 1457–1466, 2010.
- GIBSON, Stuart J. **Eigenfit: A statistical learning approach to facial composites**. 2006. Tese (Doutorado) – University of Kent.
- GREGORI, Ivan RS; SANCHES, Ivandro; THOMAZ, Carlos E. Clutch judder classification and prediction: A multivariate statistical analysis based on torque signals. **IEEE transactions on industrial electronics**, IEEE, v. 64, n. 5, p. 4287–4295, 2016.
- GRUNAU, Ruth Eckstein et al. Bedside application of the neonatal facial coding system in pain assessment of premature infants. **Pain**, Elsevier, v. 76, n. 3, p. 277–286, 1998.
- GRUNAU, Ruth VE; CRAIG, Kenneth D. Pain expression in neonates: facial action and cry. **Pain**, Elsevier, v. 28, n. 3, p. 395–410, 1987.
- GRUSS, Sascha et al. Pain intensity recognition rates via biopotential feature patterns with support vector machines. **PloS one**, Public Library of Science, v. 10, n. 10, e0140330, 2015.
- GUINSBURG. Avaliação e tratamento da dor no recém-nascido. **J Pediatr (Rio J)**, v. 75, n. 3, p. 149–60, 1999.
- GUINSBURG, Ruth; CUENCA, Maria Carmenza. A linguagem da dor no recém-nascido. **São Paulo: Sociedade Brasileira de Pediatria**. [Internet], 2010.

- GUINSBURG, Ruth et al. Reliability of two behavioral tools to assess pain in preterm neonates. **Sao Paulo Medical Journal**, SciELO Brasil, v. 121, n. 2, p. 72–76, 2003.
- HADJISTAVROPOULOS, Thomas et al. Pain assessment in elderly adults with dementia. **The Lancet Neurology**, Elsevier, v. 13, n. 12, p. 1216–1227, 2014.
- HAGER, J. C. Innovation and Intellectual Property Rights. In: EKMAN, P.; ROSENBERG, E. L (Ed.). **What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS)**. New York, NY, US: Oxford University Press, 1997. p. 58–62.
- HEIDERICH, Tatianny Marcondes; LESLIE, Ana Teresa Figueiredo Stochero; GUINSBURG. Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements. **Acta Paediatrica**, Wiley Online Library, v. 104, n. 2, e63–e69, 2015.
- HOLSTI, Liisa; GRUNAU, Ruth E. Initial validation of the behavioral indicators of infant pain (BIIP). **Pain**, Elsevier, v. 132, n. 3, p. 264–272, 2007.
- HUMMEL, P et al. Clinical reliability and validity of the N-PASS: neonatal pain, agitation and sedation scale with prolonged pain. **Journal of Perinatology**, Nature Publishing Group, v. 28, n. 1, p. 55, 2008.
- JIN, Hongliang et al. Face detection using improved LBP under Bayesian framework. In: **THIRD INTERNATIONAL CONFERENCE ON IMAGE AND GRAPHICS (ICIG'04)**. [S.l.]: IEEE, 2004. p. 306–309.
- JOHNSON, R. A.; WICHERN, D. W. Comparison of Batches. In: **Applied Multivariate Statistical Analysis**. Edição: R. A. Johnson e D. W. Wichern. 6. ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. p. 1.
- JOHNSON, R. A.; WICHERN, D. W. Comparison of Batches. In: **Applied Multivariate Statistical Analysis**. Edição: R. A. Johnson e D. W. Wichern. 6. ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. p. 8.
- JOHNSON, R. A.; WICHERN, D. W. Regression Models. In: **Applied Multivariate Statistical Analysis**. Edição: R. A. Johnson e D. W. Wichern. 6. ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. p. 430.
- KRECHEL, Susan W; BILDNER, JUDY. CRIES: a new neonatal postoperative pain measurement score. Initial testing of validity and reliability. **Pediatric Anesthesia**, Wiley Online Library, v. 5, n. 1, p. 53–61, 1995.
- LAWRENCE, Jocelyn et al. The development of a tool to assess neonatal pain. **Neonatal network: NN**, v. 12, n. 6, p. 59–66, 1993.
- LIAO, Shu; CHUNG, Albert CS. Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude. In: **8TH ASIAN CONFERENCE ON COMPUTER VISION**. Tokyo, Japan: Springer, 2007. p. 672–679.
- LU, Guanming et al. Sparse representation based facial expression classification for pain assessment in neonates. In: **NATURAL COMPUTATION, FUZZY SYSTEMS AND KNOWLEDGE DISCOVERY (ICNC-FSKD), 2016 12TH INTERNATIONAL CONFERENCE ON**. Changsha, China: IEEE, 2016. p. 1615–1619.
- LUDA et al. Genetic architecture of human pain perception. **TRENDS in Genetics**, Elsevier, v. 23, n. 12, p. 605–613, 2007.
- LYNCH, Maureen. Pain as the fifth vital sign. **Journal of Infusion Nursing**, LWW, v. 24, n. 2, p. 85–94, 2001.

- MAINTZ, JB Antoine; VIERGEVER, Max A. A survey of medical image registration. **Medical image analysis**, Elsevier, v. 2, n. 1, p. 1–36, 1998.
- MANSOR, Muhammad Naufal; REJAB, Mohd Nazri. A computational model of the infant pain impressions with Gaussian and nearest mean classifier. In: CONTROL SYSTEM, COMPUTING AND ENGINEERING (ICCSCE), 2013 IEEE INTERNATIONAL CONFERENCE ON. Penang, Malaysia: IEEE, 2013. p. 249–253.
- MANSOR, Muhammad Naufal; REJAB, Mohd Nazri. Infant pain recognition system with GLCM features and GANN under unstructured lighting condition. In: CONTROL SYSTEM, COMPUTING AND ENGINEERING (ICCSCE). Penang, Malaysia: IEEE International Conference, 2013. p. 243–248.
- MELO, Gleicia Martins de et al. Escalas de avaliação de dor em recém-nascidos: revisão integrativa. **Revista Paulista de Pediatria**, Elsevier, v. 32, n. 4, p. 395–402, 2014.
- MERKEL, Sandra I et al. The FLACC: a behavioral scale for scoring postoperative pain in young children. **Pediatric nursing**, v. 23 3, p. 293–7, 1997.
- NANNI, Loris; BRAHNAM, Sheryl; LUMINI, Alessandra. A local approach based on a Local Binary Patterns variant texture descriptor for classifying pain states. **Expert Systems with Applications**, Elsevier, v. 37, n. 12, p. 7888–7894, 2010.
- NANNI, Loris; LUMINI, Alessandra; BRAHNAM, Sheryl. Local binary patterns variants as texture descriptors for medical image analysis. **Artificial intelligence in medicine**, Elsevier, v. 49, n. 2, p. 117–125, 2010.
- O'ROURKE, Deborah. The measurement of pain in infants, children, and adolescents: from policy to practice. **Physical Therapy**, Oxford University Press, v. 84, n. 6, p. 560–570, 2004.
- OJALA, Timo; PIETIKAINEN, Matti; MAENPAA, Topi. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on pattern analysis and machine intelligence**, IEEE, v. 24, n. 7, p. 971–987, 2002.
- OKADA, Massako et al. Dor em pediatria. **Revista de Medicina**, v. 80, spe1, p. 135–156, 2001.
- PANTIC, Maja; PATRAS, Ioannis. Dynamics of facial expression: recognition of facial actions and their temporal segments from face profile image sequences. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, IEEE, v. 36, n. 2, p. 433–449, 2006.
- PANTIC, Maja; ROTHKRANTZ, Leon JM. Facial action recognition for facial expression analysis from static face images. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, IEEE, v. 34, n. 3, p. 1449–1461, 2004.
- PILLAI RIDDELL, Rebecca R; BADALI, Melanie A; CRAIG, Kenneth D. Parental judgments of infant pain: Importance of perceived cognitive abilities, behavioural cues and contextual cues. **Pain Research and Management**, Hindawi, v. 9, n. 2, p. 73–80, 2004.
- PILLAI RIDDELL, Rebecca R; CRAIG, Kenneth D. Judgments of infant pain: The impact of caregiver identity and infant age. **Journal of pediatric psychology**, Oxford University Press, v. 32, n. 5, p. 501–511, 2006.
- PILLAI RIDDELL, Rebecca R et al. Understanding caregiver judgments of infant pain: Contrasts of parents, nurses and pediatricians. **Pain Research and Management**, Hindawi, v. 13, n. 6, p. 489–496, 2008.
- REN, Jianfeng; JIANG, Xudong; YUAN, Junsong. Face and facial expressions recognition and analysis. In: REN, Jianfeng; JIANG, Xudong; YUAN, Junsong (Ed.). **Context Aware Human-Robot and Human-Agent Interaction**. 1. ed. [S.l.]: Springer, 2016. p. 3–29.

- RUECKERT, Daniel et al. Nonrigid registration using free-form deformations: application to breast MR images. **IEEE transactions on medical imaging**, IEEE, v. 18, n. 8, p. 712–721, 1999.
- SALEKIN, Sirajus. **Multimodal Assessment of Neonatal Pain Using Computer Vision**. 2017.
- SATO, João Ricardo et al. Hyperplane navigation: a method to set individual scores in fMRI group datasets. **Neuroimage**, Elsevier, v. 42, n. 4, p. 1473–1480, 2008.
- SCOPEL, Evânea; ALENCAR, Márcia; CRUZ, Roberto Moraes. Medidas de avaliação da dor. **Lecturas: Educación física y deportes**, Tulio Guterman, n. 105, p. 34, 2007.
- SERAG, Ahmed et al. Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. **NeuroImage**, Elsevier, v. 59, n. 3, p. 2255–2265, 2012.
- SHAFFER, John R et al. Genome-wide association study reveals multiple loci influencing normal human facial morphology. **PLoS genetics**, Public Library of Science, v. 12, n. 8, e1006149, 2016.
- SHANNON, Claude Elwood. A mathematical theory of communication. **ACM SIGMOBILE mobile computing and communications review**, ACM New York, NY, USA, v. 5, n. 1, p. 3–55, 2001.
- SPENCE, Kaye et al. A reliable pain assessment tool for clinical assessment in the neonatal intensive care unit. **Journal of Obstetric, Gynecologic, & Neonatal Nursing**, Wiley Online Library, v. 34, n. 1, p. 80–86, 2005.
- STEVENS, Bonnie; GIBBINS, Sharyn. Clinical utility and clinical significance in the assessment and management of pain in vulnerable infants. **Clinics in Perinatology**, v. 29, n. 3, p. 459–468, 2002.
- STEVENS, Bonnie et al. Premature Infant Pain Profile: development and initial validation. **The Clinical journal of pain**, LWW, v. 12, n. 1, p. 13–22, 1996.
- TAIGMAN, Yaniv et al. Deepface: Closing the gap to human-level performance in face verification. In: **PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION**. NW Washington, DC United States: IEEE, 2014. p. 1701–1708.
- TAN, Xiaoyang; TRIGGS, William. Enhanced local texture feature sets for face recognition under difficult lighting conditions. **IEEE transactions on image processing**, v. 19, n. 6, p. 1635–1650, 2010.
- TERUEL, Gilberto Ferreira. **Análise e reconhecimento de dor em imagens 2D frontais de recém-nascidos e termo e saudáveis**. Orientador: Carlos Eduardo Thomaz, 2018. 61 f. Dissertação (Mestrado em Engenharia Elétrica) – Centro Universitário FEI, São Bernardo do Campo, 2018.
- TERUEL, Gilberto F et al. Análise e Recon de Recém-Nascidos a Termo e Saudáveis. In: **ANAIS Estendidos do XIX Simpósio Brasileiro de Computação Aplicada à Saúde**. São Domingos, Niterói, Rio de Janeiro: SBC, 2019. p. 97–102.
- THOMAZ, Carlos E et al. A multivariate statistical analysis of the developing human brain in preterm infants. **Image and Vision Computing**, Elsevier, v. 25, n. 6, p. 981–994, 2007.
- THOMAZ, Carlos E et al. Using a maximum uncertainty LDA-based approach to classify and analyse MR brain images. In: **INTERNATIONAL CONFERENCE ON MEDICAL IMAGE**

COMPUTING AND COMPUTER-ASSISTED INTERVENTION. Saint-Malo, France: Springer, 2004. p. 291–300.

THOMAZ; GIRALDI, Gilson Antonio. A new ranking method for principal components analysis and its application to face image analysis. **Image and Vision Computing**, Elsevier, v. 28, n. 6, p. 902–913, 2010.

THOMAZ; KITANI, Edson Caoru; GILLIES, Duncan Fyfe. A maximum uncertainty LDA-based approach for limited sample size problems—with application to face recognition. **Journal of the Brazilian Computer Society**, Springer, v. 12, n. 2, p. 7–18, 2006.

THOMAZ et al. A priori-driven multivariate statistical approach to reduce dimensionality of MEG signals. **Electronics Letters**, IET, v. 49, n. 18, p. 1123–1124, 2013.

UMEYAMA, Shinji. Least-squares estimation of transformation parameters between two point patterns. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, IEEE, n. 4, p. 376–380, 1991.

VERVERIDIS, Dimitrios; KOTROPOULOS, Constantine. Fast and accurate sequential floating forward feature selection with the Bayes classifier applied to speech emotion recognition. **Signal Processing**, Elsevier, v. 88, n. 12, p. 2956–2970, 2008.

XAVIER, I et al. A photo-realistic generator of most expressive and discriminant changes in 2D face images. In: IEEE. EMERGING Security Technologies (EST), 2015 Sixth International Conference on. [S.l.: s.n.], 2015. p. 80–85.

XAVIER, I et al. Construction of a Spatio-Temporal Face Atlas: Experiments Using Down Syndrome Samples. In: IN PROCEEDINGS OF THE 29TH SIBGRAPI, CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES, WORKSHOP ON FACE PROCESSING APPLICATIONS. Sao Jose dos Campos, Sao Paulo, Brazil: IEEE, 2016. p. 1–4.

XAVIER, I et al. Construction of a Spatio-Temporal Face Atlas: Experiments Using Down Syndrome Samples. In: CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES (SIBGRAPI): WORKSHOP ON FACE PROCESSING APPLICATIONS. São José dos Campos, São Paulo, Brasil: IEEE, 2016. p. 1–4.

ZAMZMI, Ghada et al. A Review of Automated Pain Assessment in Infants: Features, Classification Tasks, and Databases. **IEEE Reviews in Biomedical Engineering**, IEEE, 2017.

ZAMZMI, Ghada et al. Neonatal Pain Expression Recognition Using Transfer Learning. **arXiv preprint arXiv:1807.01631**, 2018.

ZHANG, Shifeng et al. S3fd: Single shot scale-invariant face detector. In: PROCEEDINGS of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017. p. 192–201.

ZHAO, Guoying; PIETIKAINEN, Matti. Dynamic texture recognition using local binary patterns with an application to facial expressions. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, IEEE, n. 6, p. 915–928, 2007.

ZHI, Ruicong et al. Automatic Infants' Pain Assessment by Dynamic Facial Representation: Effects of Profile View, Gestational Age, Gender, and Race. **Journal of clinical medicine**, Multidisciplinary Digital Publishing Institute, v. 7, n. 7, p. 173, 2018.