

Neonatal Face Mosaic: An areas-of-interest segmentation method based on 2D face images

Pedro Henrique Silva Domingues
dept. de Ciência da Computação
Centro Universitário da FEI
São Paulo, Brasil
12pedro07@gmail.com

Renan Martins Mendes da Silva
dept. de Ciência da Computação
Centro Universitário da FEI
São Paulo, Brasil
renanmartins1999@hotmail.com

Ibrahim Jamil Orta
dept. de Ciência da Computação
Centro Universitário da FEI
São Paulo, Brasil
ibrahimorra459@gmail.com

Matheus Elias Cruz
dept. de Ciência da Computação
Centro Universitário da FEI
São Paulo, Brasil
matheus.elias.cruz@hotmail.com

Tatiany Marcondes Heiderich
dept. de Engenharia Elétrica
Centro Universitário da FEI
São Paulo, Brasil
tatianyms@gmail.com

Carlos Eduardo Thomaz
dept. de Engenharia Elétrica
Centro Universitario da FEI
São Paulo, Brasil
cet@fei.edu.br

Abstract—The daily life of preterm babies may be involved with long exposure to pain, causing problems in the development of the nervous system. In this context, an on-going area of research is the scientific development of image-based automatic pain detection systems based on several techniques, from anatomical measurements to artificial intelligence, they have generally two main issues: the categorization of the most relevant facial regions for identifying neonatal pain and the practical difficulty related to the presence of artifacts obstructing parts of the face. This paper proposes and implements an areas-of-interest automatic segmentation method that allows the creation of a novel dataset containing crops of neonatal faces relevant for pain classification, labelled by areas-of-interest and pain status. Moreover, we have also investigated the use of similarity matching techniques to compare each area-of-interest to the corresponding one extracted from a prototype face with no occlusion.

Index Terms—preterm infants, face dataset, face segmentation, mosaic, computer vision

I. INTRODUÇÃO

A avaliação da dor usualmente é um processo descritivo verbal simples para qualquer adulto, mas como mencionado em [1], mesmo os prematuros extremos são capazes de sentir dor e nestes casos a comunicação verbal deixa de ser uma possibilidade. Logo, visto que bebês prematuros são expostos a vários procedimentos dolorosos durante um período de rápido desenvolvimento cerebral, [2], e sabendo que as experiências de dor intensiva durante este período pode impactar no desenvolvimento dos componentes somatossensoriais e emocionais da resposta à dor na vida adulta [3], identifica-se a necessidade da criação de métodos específicos para diagnóstico da dor nestes casos.

Atualmente, existem indicadores bem estabelecidos para classificação de dor em recém-nascidos, alguns exemplos são as escalas *Bernese Pain Scale for Neonates* (BPSN), *Neonatal Infant Pain Scale* (NIPS), *Premature Infant Pain Profile-Revised* (PIPP-R) e *Neonatal Pain Agitation and Sedation Scale* (N-PASS), nas quais há medições multidimensionais,

podendo ser estas fisiológicas, como saturação do oxigênio e comportamentais, como aperto dos olhos ou agitação. Porém, quando o contexto é a medição automática da dor, alguns destes indicadores podem ser mais desafiadores e complexos de serem medidos automaticamente. Portanto, escalas unidimensionais como *Neonatal Facial Coding System* (NFCS), baseadas apenas em movimentos faciais, tornam-se mais populares neste contexto.

Por outro lado, trabalhos como [2,4,5,6] estudam formas de detecção automática de dor neonatal utilizando metodologias fundamentadas em técnicas de processamento de imagens e/ou aprendizado de máquina. Nestes trabalhos são utilizadas bases de imagens cujos recém-nascidos são vistos com baixa ou nenhuma presença de artefatos, como fixação de sondas e intubação orotraqueal (IOT), máscara de ventilação (Continuous Positive Airway Pressure - CPAP) e máscaras, obstruindo o rosto. Estes objetos são uma forma de ruído de difícil tratamento, porém em cenários reais, tais objetos são comumente encontrados, não podendo ser removidos durante extensos períodos de tempo, pois são indispensáveis para a manutenção da vida do recém-nascido.

Por este motivo, este trabalho propõe a separação da face em regiões poligonais pré-definidas e relevantes para a detecção de dor em neonatos, escolhidas com base na escala NFCS e denominadas de mosaico facial. Através deste mosaico foi criada uma base de dados para que, estudos futuros possam separar as regiões obstruídas por artefatos na imagem e realizar classificações de dor ponderadas por região. Adicionalmente, também é proposta e avaliada uma primeira forma de detecção da presença de artefatos obstruindo as regiões estudadas utilizando métricas de similaridade e uma face média construída utilizando imagens de recém-nascidos sem a presença de nenhum artefato.

II. BASE DE DADOS

Para o desenvolvimento do algoritmo foi utilizado o banco de imagens de faces de recém-nascidos da UNIFESP. O banco mencionado não é de livre acesso, porém foi autorizado para o uso durante este estudo.

As imagens foram capturadas antes, durante e depois do procedimento de punção, com intervalos de 45 a 50 segundos e em um grupo de 30 recém-nascidos, dos quais seis foram filmados na UTI neonatal e 24 no alojamento conjunto. O total de imagens foi 360, sendo 238 categorizadas por especialistas como 'com dor' e 122 como 'sem dor'. As imagens são de baixa resolução (em média 450x230 px) e se encontram no formato bmp [2].

Além das imagens já presentes na base de dados, foram separadas mais 10 imagens de recém-nascidos, porém com a presença de artefatos obstruindo parcialmente regiões da face, utilizadas para testes.

III. MOSAICO

As regiões de interesse utilizadas no mosaico foram escolhidas com base na escala NFCS, uma métrica utilizada por profissionais de saúde para identificar a dor em neonatos através de expressões faciais das regiões: sobrancelhas, testa, olhos, sulco nasolabial, nariz, bochecha e boca.

Cada região contribui com 1 ponto em uma escala de 0 a 8, tal que, para uma pontuação maior ou igual a 3, interpreta-se como presença de dor.

No desenvolvimento da pesquisa, decidiu-se separar o espaço entre os olhos, da testa, formando duas regiões independentes. Todas as partes do mosaico possíveis foram subdivididas entre esquerda e direita, visando enriquecer o conteúdo do *dataset* criado neste trabalho.

O mosaico foi gerado realizando a ligação entre os pontos fiduciais da face (Figura 1). Os pontos que descrevem cada região são:

- 1) Sobrancelha direita: 43 a 51.
- 2) Sobrancelha esquerda: 97 a 105
- 3) Olho direito: 33 ao 37, 39, 43 a 47, 75
- 4) Olho esquerdo: 81, 87, 89, 90, 91, 93, 97 a 101
- 5) Nariz: 72, 75 a 85
- 6) Entre olhos: 39, 46, 49 a 51, 72, 75, 81, 89, 97, 102 a 104
- 7) Boca: 52, 53, 55, 56, 58, 59, 61, 63, 64, 67, 68
- 8) Sulco Nasolabial Esquerdo: 20 a 22, 61, 82, 83, 86".
- 9) Sulco Nasolabial Direito: 4 a 6, 52, 76, 77, 86'.
- 10) Bochecha esquerda: 18 a 20, 30 a 32, 82, 86".
- 11) Bochecha direita: 2 a 4, 14 a 16, 76, 86'
- 12) Testa: 1, 9' a 14', 25' a 30', 17.

IV. METODOLOGIA

A metodologia deste trabalho foi dividida em duas etapas, a primeira refere-se a identificação, transformação e extração das regiões de interesse através de um *pipeline*. A segunda consiste na montagem de uma face média dos recém-nascidos e uso de métricas de similaridade para verificar a presença de artefatos.

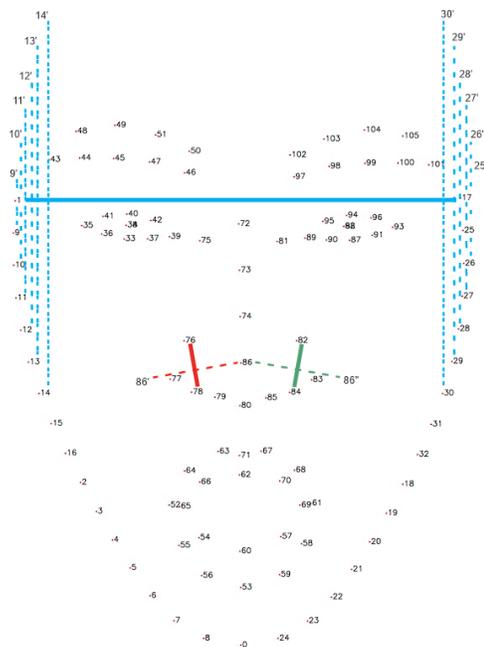


Fig. 1. Novos pontos criados a partir dos já existentes no conjunto após a regressão. Adaptado de [7].

A. Pipeline

A imagem do recém-nascido passa pelas seguintes etapas em ordem:

- Uma rede neural *RetinaFace* detecta a face do recém-nascido, identificando um *bounding box* e cinco pontos chave, dois no centro dos olhos, um para a ponta do nariz e dois nas laterais do lábio (Figura 2 B). Este resultado passa por um regressor para que 106 novos pontos sejam encontrados (Figura 2 C) [7];
- A face é recortada a partir da *bounding box* encontrada, a imagem resultante é redimensionada para ter dimensão 68x68 pixels (valor escolhido em função da menor face encontrada no *dataset* da UNIFESP utilizado) utilizando uma borda preta para que a face mantenha suas proporções e não seja distorcida no processo (Figura 2 D);
- As coordenadas dos pontos chave são então transformadas para que possam ser utilizadas na nova imagem gerada;
- Uma transformação rígida faz com que a reta dos olhos do recém-nascido (pontos 35 e 93) alinhe-se com a reta $y = 22$ (Figura 2 E). A seguir, os pontos são novamente transformados para que acomodem a nova posição da face;
- Um grupo novo de pontos é gerado espelhando o contorno superior da face (pontos 9 a 14 e 25 a 30) na reta formada por 1 e 17 (conforme indicado em azul na Figura 1). Desta forma é possível aproximar a região da testa do recém-nascido.
- Outros dois pontos são criados a partir da mesma técnica,

porém espelhando o ponto 86 nas retas definidas por [76, 78] e [82, 84] (indicado na Figura 1 em vermelho e verde). Estes novos pontos são utilizados para auxiliar a delimitar a região dos sulco nasolabiais e bochechas esquerda e direita;

- Um mosaico é criado, separando todas as regiões de interesse a serem estudadas (Figura 2 F): Testa, olhos, sobrancelhas, região entre olhos, nariz, bochechas, sulcos nasolabiais e boca. Visando adicionar uma margem de erro, todas as regiões são dilatadas em 10% (Figura 2 G);
- Por fim, são criadas máscaras binárias para cada região (Figura 2 H), as quais são aplicadas individualmente, gerando doze imagens, uma para cada região de interesse (Figura 2 I).

O *Pipeline* foi utilizado em todas as imagens da base da UNIFESP, gerando a nova base com os recortes categorizados como dor e não dor.

B. Similaridade

Para esta etapa, o dataset da UNIFESP foi dividido em duas partes, uma com 40 imagens, sendo 20 de recém-nascidos com dor e outras 20 sem dor, utilizadas para testes e outra parte com todas as imagens restantes (218 com dor e 102 sem dor) utilizadas para a criação da face média após a transformação rígida aplicada no *pipeline* (Figura 3). A face média gerada utilizou a mediana da intensidade de cada pixel das imagens 68x68.

Para obter o valor de similaridade, a imagem do recém nascido passa pelo *pipeline* e a mascara de cada região é aplicada também à face média. Após isso, os pixels da região de interesse das duas imagens foram comparados utilizando três métricas escolhidas pelos autores, sendo estas: erro médio quadrático (MSE), coeficiente de correlação de Pearson (PCC) e informação mútua (MI).

V. RESULTADOS E DISCUSSÃO

A. Base de Dados

A base de dados gerada como resultado deste trabalho foi resultado direto do processamento da base da UNIFESP pelo *pipeline* descrito na seção IV-A.

A estrutura estabelecida seguiu o formato demonstrado na Figura 4.

Os recortes do *dataset* criado estão em formato jpg, coloridos, possuem resolução 68x68 pixels e ficam dentro do diretório *Regiões*. As máscaras se encontram no mesmo formato e resolução, porém são salvas com apenas 1 canal de cor com valores 255 para a região de interesse e 0 para as demais.

Qualquer imagem de região pode ser mapeada para a relativa máscara ou imagem de origem no *dataset* da UNIFESP utilizando o nome do arquivo.

B. Verificação de artefatos

Utilizando as 40 imagens para teste mencionadas na seção IV-B, foram gerados os resultados apresentados na tabela I

TABLE I
MÉTRICAS POR REGIÃO

Região	MI		MSE		PCC	
	μ	σ	μ	σ	μ	σ
Testa	0.65	0.14	1987	1493	0.05701	0.02105
Sobrancelha esquerda	0.32	0.16	1126	2137	0.00340	0.00401
Sobrancelha direita	0.31	0.13	1419	1600	0.00228	0.00257
Entre olhos	0.39	0.15	1142	1610	0.00670	0.00430
Olho esquerdo	0.34	0.14	1140	1993	0.00596	0.00390
Olho direito	0.36	0.15	1421	1391	0.00689	0.00441
Nariz	0.31	0.11	1331	1286	0.00302	0.00390
Bochecha esquerda	0.50	0.15	1937	2500	0.01004	0.01622
Bochecha direita	0.38	0.22	1939	1562	0.00946	0.02448
Boca	0.43	0.14	1573	1258	0.00869	0.00994
Sulco nasolabial esquerdo	0.35	0.15	1616	1693	0.00903	0.01677
Sulco nasolabial direito	0.40	0.18	2184	1695	0.00618	0.02292

para as métricas escolhidas comparando regiões sem artefatos com as mesmas na face média.

Adicionalmente, foram comparadas algumas regiões obstruídas por artefatos para validar a capacidade de separação destas por meio de cada métrica. A Figura 5 apresenta o resultado da aplicação do *pipeline* seguido dos valores de similaridade obtidos para três regiões distintas. O valor de cada métrica é apresentado como uma linha tracejada vermelha e os histogramas representam o resultado da mesma métrica e região, extraído do conjunto de imagens de teste utilizado na elaboração da Tabela I

Observando os resultados obtidos, determinou-se que para as métricas escolhidas, a informação mútua (MI) não foi capaz de validar a presença de artefato, o coeficiente de correlação de Pearson (PCC) foi capaz de validar em parte dos casos, porém com uma variação muito baixa de similaridade, por fim, o erro médio quadrático (MSE) se mostrou o mais promissor na validação da presença de artefatos obstruindo a região.

VI. CONCLUSÃO

Este artigo propôs um *pipeline* capaz de receber como entrada a imagem de um recém-nascido e retornar como saída um conjunto de máscaras e recortes de regiões da face relevantes para a classificação de dor. O conjunto destas regiões foi nomeado como mosaico facial e a partir dele foi criada uma base de dados derivada da base de imagens de recém-nascidos da UNIFESP.

Complementarmente, foi proposto o uso de métricas de similaridade aplicadas a base de dados gerada, comparando as regiões do mosaico e as mesmas regiões retiradas de uma face média, visando identificar a existência de obstrução por artefatos. Das métricas estudadas, o erro médio quadrático demonstrou-se a mais promissora para este fim.

Por fim, espera-se que em trabalhos futuros possamos testar mais métricas e técnicas para detectar a presença de artefatos

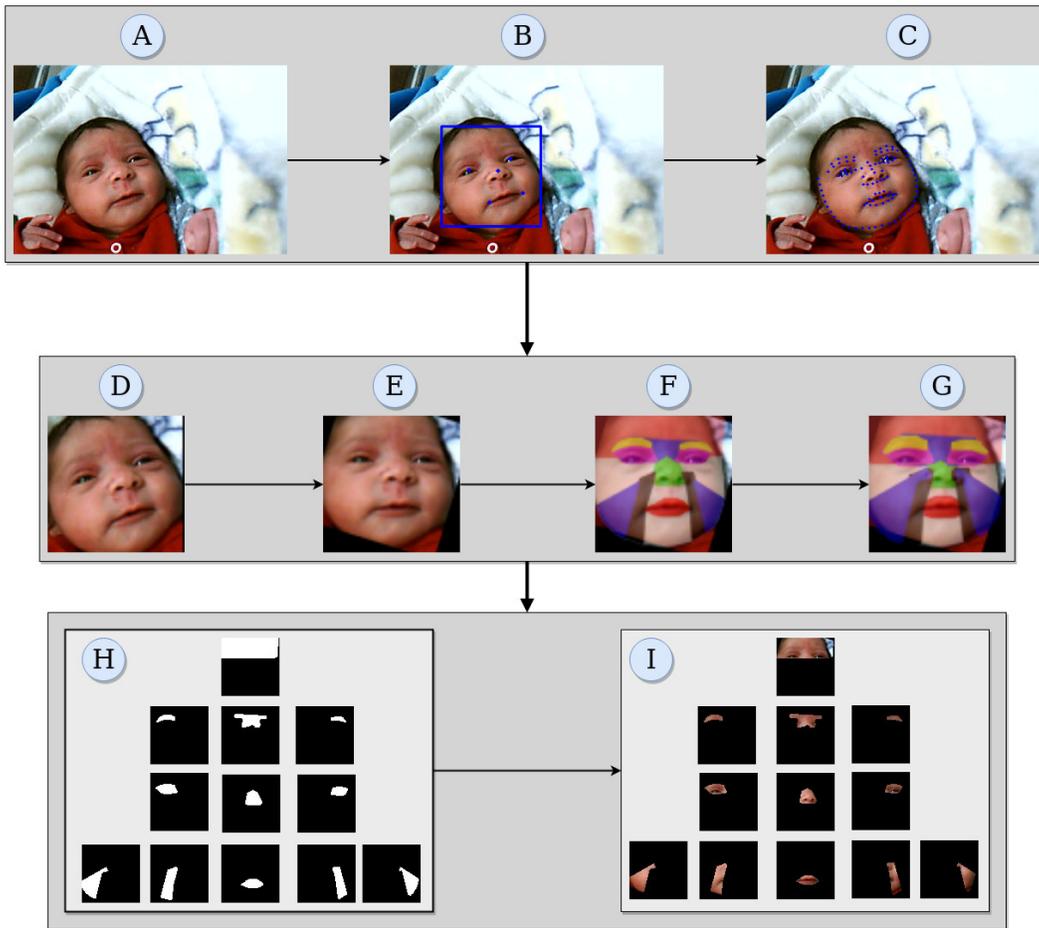


Fig. 2. A. Imagem de entrada; B. Detecção com RetinaFace; C. Regressão dos pontos fiduciais; D. Recorte da face; E. Transformação Rígida; F. Mosaico; G. Dilatação dos polígonos do mosaico em 10%; H. Conjunto de máscaras binárias geradas; I. Recortes gerados.

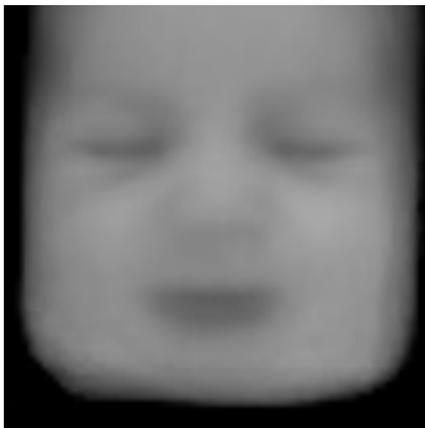


Fig. 3. Face média para os recém nascidos da base da UNIFESP 68x68px ampliada.

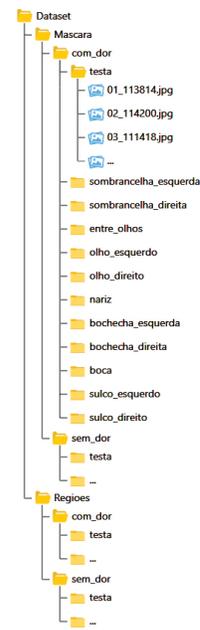


Fig. 4. Estrutura do dataset.

em cada região e utilizar a base de mosaicos criada para a tarefa de detecção de dor por região.

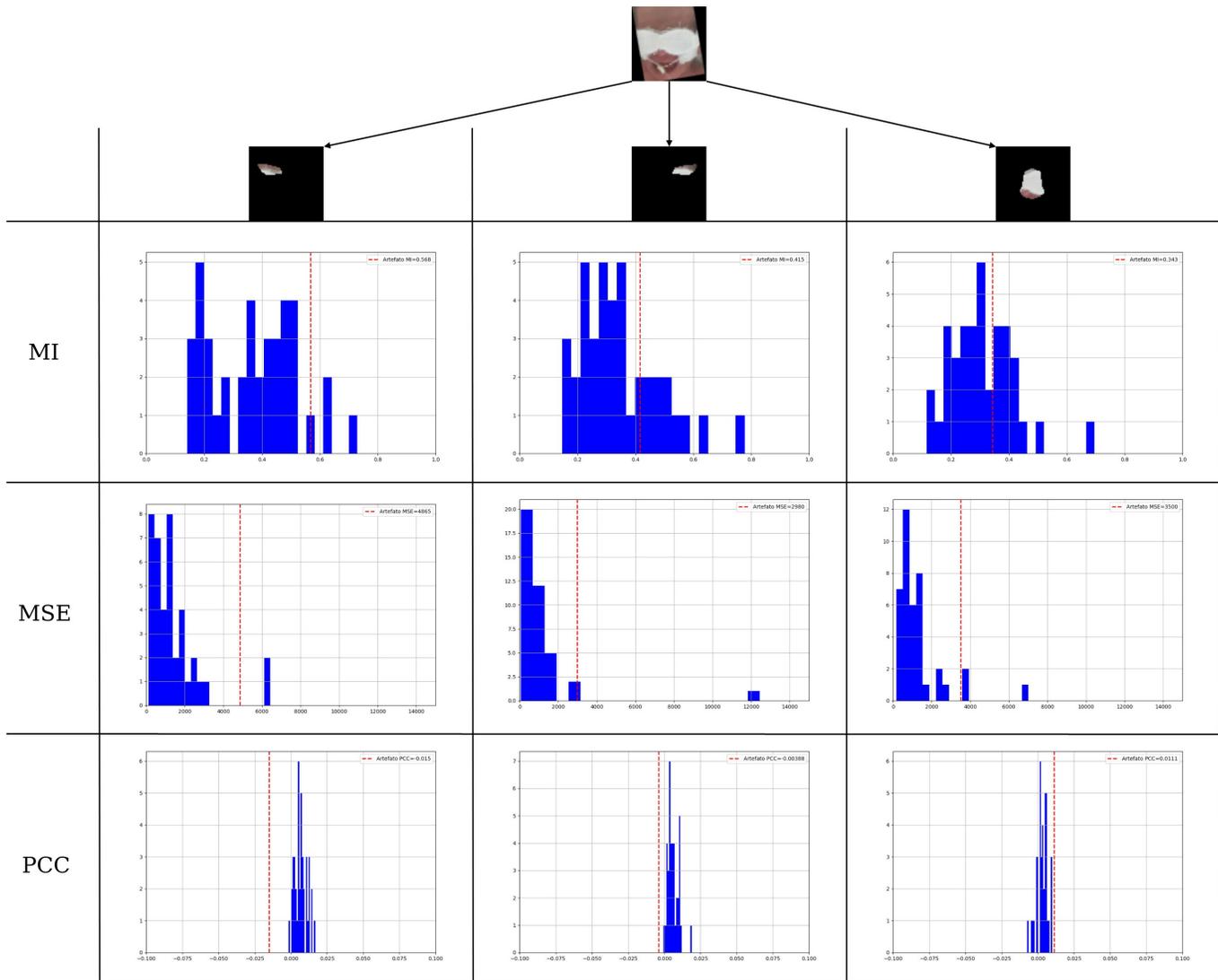


Fig. 5. Similaridade obtida para regiões com artefatos comparada com o histograma das mesmas regiões sem artefatos.

REFERENCES

- [1] R. d. C. X. Balda, R. Guinsburg, "Evaluation and treatment of pain in the neonatal period," *Revista Pediátrica-Publicação Oficial da Sociedade Brasileira de Pediatria*, v. 9, n. 1, p. 43–52, 2019, doi: 10.25060/residpediatr-2019.v9n1-13.
- [2] T. M. Heiderich, A. T. F. S. Leslie, R. 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, vol. 104, n. 2, pp. e63–e69, 2015.
- [3] S. M. Walker, "Long-term effects of neonatal pain" In: *ELSEVIER.Seminars in Fetal and Neonatal Medicine*. [S.l.], 2019. v. 24, n. 4, p. 101005.
- [4] G. F. Teruel, C. E. Thomaz, "Análise e reconhecimento de dor em imagens 2d frontais de recém-nascidos a termo e saudáveis," In: *SBC. Anais Estendidos do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*. [S.l.], 2019. p. 97–102.
- [5] G. Zamzmi, S. Salekin, D. Goldgof, T. Ho, Y. Sun, "Pain assessment from facial expression: Neonatal convolutional neural network (n-cnn)." In: *IEEE.2019 International Joint Conference on Neural Networks(IJCNN)*. [S.l.], 2019. p. 1–7.
- [6] L. Buzuti, "Avaliação de dor em expressão facial neonatal por meio de redes neurais profundas". Centro Universitário FEI, São Bernardo do Campo, 2020.
- [7] J. Deng, J. Guo, E. Ververas, I. Kotsia, S. Zafeiriou, "Retinaface: Single-shot multi-level face localisation in the wild". In: *IEEE.2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.], 2020. p. 5202–5211.