



REVIEW ARTICLE

Face-based automatic pain assessment: challenges and perspectives in neonatal intensive care units



Tatiany M. Heiderich ^{a,*}, Lucas P. Carlini ^a, Lucas F. Buzuti ^a, Rita de C.X. Balda ^b, Marina C.M. Barros ^b, Ruth Guinsburg ^b, Carlos E. Thomaz ^a

^a Centro Universitário da Fundação Educacional Inaciana (FEI), São Bernardo do Campo, SP, Brazil

^b Universidade Federal de São Paulo (UNIFESP), São Paulo, SP, Brazil

Received 8 March 2023; accepted 22 May 2023

Available online 15 June 2023

KEYWORDS

Pain measurement;
Newborn;
Artificial intelligence;
Computer systems;
Software;
Automated facial
recognition

Abstract

Objective: To describe the challenges and perspectives of the automation of pain assessment in the Neonatal Intensive Care Unit.

Data sources: A search for scientific articles published in the last 10 years on automated neonatal pain assessment was conducted in the main Databases of the Health Area and Engineering Journal Portals, using the descriptors: Pain Measurement, Newborn, Artificial Intelligence, Computer Systems, Software, Automated Facial Recognition.

Summary of findings: Fifteen articles were selected and allowed a broad reflection on first, the literature search did not return the various automatic methods that exist to date, and those that exist are not effective enough to replace the human eye; second, computational methods are not yet able to automatically detect pain on partially covered faces and need to be tested during the natural movement of the neonate and with different light intensities; third, for research to advance in this area, databases are needed with more neonatal facial images available for the study of computational methods.

Conclusion: There is still a gap between computational methods developed for automated neonatal pain assessment and a practical application that can be used at the bedside in real-time, that is sensitive, specific, and with good accuracy. The studies reviewed described limitations that could be minimized with the development of a tool that identifies pain by analyzing only free facial regions, and the creation and feasibility of a synthetic database of neonatal facial images that is freely available to researchers.

© 2023 Sociedade Brasileira de Pediatria. Published by Elsevier Editora Ltda. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Corresponding author.

E-mail: tatianymh@fei.edu.br (T.M. Heiderich).

Introduction

Facial expression analysis is a non-invasive method for pain assessment in premature and full-term newborns frequently used in Neonatal Intensive Care Units (NICU) for pain diagnosis.¹ When newborns experience a painful sensation, the facial features observed are brow bulge, eye squeeze, nasolabial furrow, open lips, stretched mouth (vertical or horizontal), lip purse, taut tongue, and chin quiver.¹ These features are present in more than 90% of neonates undergoing painful stimuli, and 95–98% of term newborns undergoing acute painful procedures exhibit at least the first three facial movements.¹ The same characteristics are absent when these patients suffer an unpleasant but not painful stimulus.^{1,2}

Several pain scales have been developed for the assessment of neonatal pain. These scales contemplate the facial expression analysis and are commonly used in the NICU, as follows: Premature Infant Pain Profile (PIPP) and PIPP Revisited (PIPP-R);^{3,4} Neonatal Pain, Agitation, and Sedation Scale (N-PASS Scale);⁵ Neonatal Facial Coding System (NFCSS);¹ Echelle de la Douleur Inconfort Nouveau-né (EDIN Scale);⁶ Crying, requires increased oxygen administration, increased vital signs, Expression, Sleeplessness Scale (CRIES Scale);⁷ COMFORT neo Scale;⁸ COVERS Neonatal Pain Scale;⁹ PAIN Assessment in Neonates Scale (PAIN Scale);¹⁰ Neonatal Infant Pain Scale (NIPS Scale).¹¹ In clinical practice, it is necessary to evaluate the scope of application of the different scales to flexibly choose the appropriate scale.¹²

Due to the wide spectrum of available different scoring methods and instruments for the diagnosis of neonatal pain, health professionals need an extensive set of skills and knowledge to conduct this task.¹³ Although some health professionals recognize the occurrence of pain in the neonatal population, the facial assessment of pain is still performed empirically in real clinical situations. One way to minimize this problem would be the use of a computational tool capable of identifying pain in critically ill newborns by evaluating facial expressions automatically and in real time.

In the last years, computational methods have been developed to detect the painful phenomenon automatically,^{14–26} to help health professionals to monitor the presence of pain and identify the need for therapeutic intervention. Even with the advancement of technology, these studies, all of them related to automatic neonatal pain assessment, have not addressed practical difficulties in identifying pain in newborns who remain with devices attached to their faces. This gap is due to the difficulty in assessing facial expression in a neonate whose face is partially covered by devices, such as enteral/gastric tube fixation, orotracheal intubation fixation, and phototherapy goggles. These problems highlight the need to develop neonatal facial movement detection techniques.

In this context, this study aims to describe the challenges and perspectives of the process of automation of neonatal pain assessment in the NICU. Specifically, the authors propose to discuss: (i) the availability of access to the literature on computational methods for automatic neonatal pain assessment (when the literature review is done in the main Databases of the Health and Engineering Areas); (ii) the computational methods available so far for the automatic evaluation of neonatal pain; (iii) the difficulty of evaluating a face that is partially covered by assistive devices; (iv) the

reduced number of databases of neonatal facial images that hinder the advance in research; (v) the perspectives for pain evaluation through the analysis of segmented facial regions.

The authors believe that this critical and up-to-date review is necessary for both the medical staff, who aim to choose an automatic method to assist in pain assessment over a continuous period; and for software engineers, who seek a starting point for further research related to the real needs of neonates in Intensive Care Units.

Method

In order to describe the challenges related to finding the available scientific literature that enables evidence-based clinical practice, the authors searched for scientific articles published in the last 10 years on the automatic assessment of neonatal pain.

The authors have searched the main Health Area Databases²⁷ and Engineering Journal Portals²⁸ (VHL - Virtual Health Library; Portal CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Embase/Elsevier; Lilacs; Medline; Pubmed, Scielo; DOAJ - Directory of Open Access Journals; IEEE Xplore - Institute of Electrical and Electronics Engineers), Semantic Scholar Database, and the arXiv Free Distribution Service.

The literature search took place in August and September 2022, using the Health Descriptors (DeCS structured vocabulary found on the Virtual Health Library site - VHL),²⁹ in English: Pain Measurement, Newborn, Artificial Intelligence, Computer Systems, Software, Automated Facial Recognition, with the Boolean operator AND.

The search for scientific articles included literature published in the last ten years on facial assessment of neonatal pain, selected from a search with the following associated descriptors: Pain Measurement and Newborn and Artificial Intelligence; Pain Measurement and Newborn and Computer Systems; Pain Measurement and Newborn and Software; Pain Measurement and Newborn and Automated Facial Recognition.

Review articles, manuscripts that did not address automated facial assessment for neonatal pain diagnosis, and duplicates were excluded.

The results were descriptive and aimed to identify the computational methods that have advanced in automating the facial assessment of neonatal pain in recent years. For this, data related to the methodology applied in each study were tabulated, as follows: the pain scale on which each study was based for the diagnosis of pain; the database and the sample used in the research; the facial regions that participated in the pain assessment, and diagnosis automation process; sensitivity and specificity in the result of each research; as well as the limitations of each study and the future perspectives of each author.

Challenging issues

Availability of literature related to automatic pain assessment in newborns

In this research, the authors identified relevant studies for the process of automation of neonatal pain assessment.

Table 1 Number of publications found in the last ten years.

Descriptors	Pain Measurement	Pain Measurement AND Newborn	Search 1 = Pain Measurement AND Newborn AND Artificial Intelligence	Search 2 = Pain Measurement AND Newborn AND Computer Systems	Search 3 = Pain Measurement AND Newborn AND Software AND Automated Facial Recognition	Duplicated in the four searches of the same Base	Total publications in the last ten years	Non-inclusion criteria	INCLUSION/ [REFERENCE]
Databases									
arXiv	92	1	0	1	0	0	1	1	0
DOAJ	2.786	14	0	1	0	0	3	3	0
Embase/ Elsevier	189.086	7.825	264	1.212	199	0	3.692	3.692	0
IEEE Xplore	778	13	3	4	3	4	6	3	3/30-32
Lilacs	1.138	61	0	0	0	0	6	6	0
Medline	55.144	869	0	2	0	0	15	14	1/14
Portal CAPES	50.045	767	5	5	1	0	16	15	1/30
Pubmed	51.779	862	4	2	1	0	21	18	3/14,30,33
Scielo	424	18	0	0	0	0	0	0	0
Semantic Scholar	2.900.000	32.500	56	378	1.050	6	1645	1636	9/23,3134-40
VHL	56.751	946	0	2	19	1	21	19	2/14,30
TOTAL	410.923	43.876	332	1.607	3.126	11	5.426	5.407	19

Note: DOAJ, Directory of Open Access Journals; IEEE Xplore, Institute of Electrical and Electronics Engineers; Lilacs, Literatura Latino-americana e do Caribe em Ciências da Saúde; Medline, Medical Literature Analysis and Retrieval System Online; Portal CAPES, Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; Scielo, Scientific Electronic Library Online; VHL, Virtual Health Library.

When performing the literature search in 11 databases (Table 1), 19 articles were found, six of them in more than one database. Two studies by Zamzmi[20,26] were added because they were cited in Grifantini's report,³⁰ totaling 15 articles[14,20,23,26,30-40] selected for review (Table 2).

It is worth noting that these 2 added articles were not found using the selected descriptors. This dissonance showed us one of the challenges in the literature search: depending on the keywords used for the search, researchers may not find relevant studies on the topic.

One way to maximize the search for scientific documents would be to systematize the search process in all databases. For this, using words common to all search systems would be interesting. A warning to researchers would be to use keywords in their publications that address both concepts related to the Health and Engineering Areas.

Table 3 shows the methods used in each study, the pain scale on which each study was based for the pain diagnosis, the database, the facial regions needed for face detection and diagnosis of neonatal pain, and the diagnostic accuracy of each method. As for the method, it was possible to observe that each study used a different method for pain detection. Interestingly, these methods did not observe the need to detect some facial regions for pain diagnoses, such as cheeks, nose, and chin. Even so, the studies were not shown to be effective to the point of being used in clinical practice because the methods developed so far have not been tested at the bedside. Each study's limitations and future perspectives are summarized and reported in Table 4.

Computational methods available for automatic neonatal pain assessment

In 2006, a pioneering study was conducted to classify facial expressions of pain. The authors applied three feature extraction techniques: principal component analysis, linear discriminant analysis, and support vector machine. The face image dataset was captured during cradling (a disturbance that can provoke crying that is not in response to pain), an air stimulus on the nose, and friction on the external lateral surface of the heel. The model based on the support vector machine achieved the best performance: pain versus non-pain 88%; pain versus rest 94.6%; pain versus cry 80%; pain versus air puff 83%; and pain versus friction 93%. The results of this study suggested that the application of facial classification techniques in pain assessment and management was becoming a promising area of investigation.⁴¹

In 2008, one of the first attempts to automate facial expression assessment of neonatal pain was performed in a study developed to compare the distance of specific facial points. However, the user manually detected each facial point, so the method was of great interest for clinical research, but not for clinical use.⁴²

In 2015, Heiderich et al. developed software to assess neonatal pain. This software was capable of automatically capturing facial images, comparing corresponding facial landmarks, and diagnosing pain presence. The software demonstrated 85% sensitivity and 100% specificity in the detection of neutral facial expressions, and 100% sensitivity and specificity in the detection of pain during painful procedures.¹⁴

Table 2 Summary of the 15 articles found.

Reference	Authors' names/ Country	Title	Study Type	Objective
¹⁴	Heiderich, T et al. (2015)/ Brasil	Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements.	Software development for neonatal pain assessment.	Develop and validate computer software to monitor neonatal facial movements of pain in real-time.
²³	Carlini, L et al. (2021)/Brasil	A Convolutional Neural Network-based Mobile Application to Bedside Neonatal Pain Assessment.	A mobile application for smartphones for neonatal pain assessment.	Propose and implement a mobile application for smartphones that uses Artificial Intelligence (AI) techniques to automatically identify the facial expression of pain in neonates, presenting feasibility in real clinical situations.
³⁰	Grifantini, C (2020)/ USA	Detecting Faces, Saving Lives.	Report: Discuss how facial recognition software is changing health care.	Report on research using facial recognition technology, with machine learning algorithms and neural networks, and could be incorporated into hospitals to reduce pain and suffering and save lives.
³⁰ cited ²⁰	First citation in Grifantini, C (2020) Zamzmi, G et al. (2019)/USA	Convolutional Neural Networks for Neonatal Pain Assessment	Investigate the use of Convolutional Neural Networks for assessing neonatal pain.	Investigate the use of a novel lightweight neonatal convolutional neural network as well as other popular convolutional neural network architectures for assessing neonatal pain.
³⁰ cited ²⁶	Second citation in Grifantini, C (2020) Zamzmi, G et al. (2022)/USA	A Comprehensive and Context-Sensitive Neonatal Pain Assessment Using Computer Vision	Present an automated system for neonatal pain assessment.	Present a pain assessment system that utilizes facial expressions along with crying sounds, body movement, and vital sign changes.
³¹	Egede, J et al. (2019)/United Kingdom	Automatic Neonatal Pain Estimation: An Acute Pain in Neonates Database.	Present an automated system for neonatal pain assessment.	Present a system for neonatal pain assessment which encodes pain indicative-features
³²	Martinez-B, A et al. (2014)/Spain	An Autonomous System to Assess, Display and Communicate the Pain Level in Newborns.	Present an automated system for neonatal pain assessment - Web application.	Present a system that automatically analyses the pain or discomfort levels of newborns.

Table 2 (Continued)

Reference	Authors' names/ Country	Title	Study Type	Objective
³³	Roué, J et al. (2021)/ France	Using sensor-fusion and machine-learning algorithms to assess acute pain in non-verbal infants: a study protocol.	The study protocol, Clinical Trials, and Prospective observational study.	Identify the specific signals and patterns from each facial sensor that correlate with the pain stimulus.
³⁴	Cheng, X et al. (2022) /China	Artificial Intelligence Based Pain Assessment Technology in Clinical Application of Real-World Neonatal Blood Sampling.	Prospective study - The client-server model to run on the mobile.	Analyze the consistency of the NPA results from a self-developed automated NPA system and nurses' on-site NPAs (OS-NPAs).
³⁵	Domingues, P et al. (2021)/Brasil	Neonatal Face Mosaic: An areas-of-interest segmentation method based on 2D face images.	Create a facial mosaic to aid in the facial assessment of neonatal pain.	Propose the separation of the face into predefined polygonal regions relevant to pain detection in neonates.
³⁶	Han, J et al. (2012)/ Netherlands	Neonatal Monitoring Based on Facial Expression Analysis.	Development of a system to analyze various facial regions.	Design a prototype of an automated video monitoring system for detecting discomfort in newborns by analyzing their facial expression.
³⁷	Mansor, M et al. (2014)/Malaysia	Infant Pain Detection with Homomorphic Filter and Fuzzy k-NN Classifier.	Present an automated system for neonatal pain assessment.	Evaluate the performance of illumination levels for infant pain classification.
³⁸	Parodi, E et al. (2017)/Italy	Automated Newborn Pain Assessment Framework Using Computer Vision Techniques.	Proposed algorithm for neonatal pain assessment.	Propose a computerized tool for neonatal pain evaluation based on patients' facial expressions.
³⁹	Wang, Y et al. (2022)/China	Full-convolution Siamese network algorithm under deep learning used in tracking of facial video image in newborns.	Explore a new tracking network for neonatal pain assessment.	Explore the full-convolution Siamese network in neonatal facial video image tracking application.
⁴⁰	Dosso, Y et al. (2022)/Canada	NICUface: Robust Neonatal Face Detection in Complex NICU Scenes	Creation of robust NICU-face detectors.	Create two neonatal face detection models (NICUface) by finetuning the most performant pre-trained face detection models on exceptionally challenging NICU scenes.

Note: AI, Artificial intelligence; Fuzzy k-NN, Fuzzy K-Nearest Neighbor; NICU, Neonatal Intensive Care Unit; NPA, Neonatal Pain Assessment; OS-NPAs, on-site NPAs.

In 2016, a study based on Machine Learning¹⁵ proposed an automated multimodal approach that used a combination of behavioral and physiological indicators to assess newborn pain. Pain recognition yielded 88%, 85%, and 82% overall

accuracy using solely facial expression, body movement, and vital signs, respectively. The combination of facial expression, body movement, and changes in vital signs (i.e., the multimodal approach) achieved 95% overall accuracy.

Table 3 Result of the literature search.

Reference	Method/Scale which was based	Database used (Sample)	Discrimination of facial regions	Sensitivity and specificity
¹⁴	Developed in the Delphi environment, based on image recognition of pain-related facial actions. /Scale: NFCS	Own Base - UNIFESP University Hospital (30 newborns between 35 and 41 weeks of gestational age).	Bulging brow; narrowing of the lid slit; deepening of the nasolabial furrow; open lips; mouth stretching.	The software exhibited 85% sensitivity and 100% specificity in detecting neutral facial expressions in their resting state and 100% sensitivity and specificity in detecting procedural pain in neonates.
²³	A computational model with face detection, data augmentation, and classification model with transfer learning to a CNN architecture pre-trained by adding fully connected layers specifically trained with neonatal face images. The application was developed for the Android operating system using the Android Studio IDE. /Scale: NFCS	UNIFESP University Hospital (30 newborns between 35 and 41 weeks of gestational age) and Infant COPE (26 Caucasian neonates).	Not applicable.	This model achieved 93.07% accuracy, 0.9431 F1 Score, and 0.9254 AUC.
³⁰	Report /Scale: Not applicable.	Not applicable.	Not applicable.	Not applicable.
³⁰ cited ²⁰	Compares the use of a novel Convolutional Neural Networks Neonatal along with others (ResNet50 and VGG-16) for pain assessment application. /Scale: NIPS	Infant COPE (26 Caucasian neonates) and NPAD (31 neonates between 32 and 40 weeks of gestational age).	Not applicable.	Assessing neonatal pain using LBP features achieved 86.8% average accuracy; Assessing neonatal pain using HOG features with Support vector machines achieved 81.29% average accuracy; Proposed N-CNN, which extracts features directly from the images, achieved state-of-the-art results and outperformed ResNet, VGG-16, as well as hand-crafted descriptors.
³⁰ cited ²⁶	Existing static methods have been divided into two categories: handcrafted-representation-based methods and deep-representation-based	Infant COPE (26 Caucasian neonates).	Not applicable.	The system achieved 95.56% accuracy using decision fusion of different pain responses that were recorded in a challenging clinical environment.

Table 3 (Continued)

Reference	Method/Scale which was based	Database used (Sample)	Discrimination of facial regions	Sensitivity and specificity
31	methods. /Scale: NIPS Uses handcrafted algorithms and deep-learned features. /Scale: NIPS and NFCS	Own Base - APN-db (213 newborns between 26 and 41 weeks of gestational age).	Brow bulge, eye squeeze, nasolabial furrow, open lips, stretch mouth (vertical), stretch mouth (horizontal), lip purse, taut tongue, chin quiver.	The system performs well with an RMSE of 1.94 compared to human error of 1.65 on the same dataset, demonstrating its potential application to newborn health care.
32	The behavioral parameters related to movement and expression are measured using computer vision techniques. /Scale: NIPS, BPSN, DAN, NFCS, PIPP and CRIES	Not reported.	Head movement, expression of pain, frowning, lips movement, eyes open/closed, cheek frowning.	Not reported.
33	Uses facial electromyography to record facial muscle activity-related infant pain. /Scale: N-PASS, PIPPR, NFCS, FLACC and VAS	Own Base (The painful procedures will be a minimum of 60 newborns and infants averaging 6 months).	Forehead, cheek, eyebrow puffing, eye pinch, and nasolabial sulcus.	Tests will be performed in further studies.
34	It was implemented with the client-server model and designed to run on the mobile nursing personal digital assistant device. /Scale: NIPS	Own Base (232 newborns with a mean gestational age of 33.93 ± 4.77 weeks).	Frown, eye squeezing, nasolabial fold deepening, mouth stretching, and tongue tightening.	The accuracies of the NIPS pain score and pain grade given by the automated NPA system were 88.79% and 95.25%, with kappa values of 0.92 and 0.90 ($p < 0.001$), respectively.
35	Identifying, transforming, and extracting the regions of interest from the face, assembling an average face of the newborns, and using similarity metrics to check for artifacts. /Scale: NFCS	UNIFESP University Hospital (30 newborns between 35 and 41 weeks of gestational age).	Eyebrows, eyes, nose, the region between the eyes, mouth, nasolabial folds, cheeks, and forehead.	Not reported. However, all images could be mapped and segmented by region.
36	The system consists of several algorithmic components, ranging from face detection, determination of the region of interest, and facial feature extraction to behavior stage classification. /Scale: Unmentioned	Own Base (newborn with different conditions)	Eyes, eyebrows, and mouth.	The algorithm can operate with approximately 88% accuracy.
37			Not applicable.	

Table 3 (Continued)

Reference	Method/Scale which was based	Database used (Sample)	Discrimination of facial regions	Sensitivity and specificity
	The Local Binary Pattern features are computed in the Fuzzy k-NN classifier employed to classify newborn pain. /Scale: Unmentioned	Infant COPE (26 Caucasian neonates)		Using the HOMO method, the sensitivity is 96.667%. Specificity ranged from 96.5% to 97.2% and accuracy ranged from 93.3% to 97.2% depending on the illumination. The fastest time consumption was obtained by Conventional Validation under 100 illumination levels with 0.065s.
38	Facial features were extracted through different image processing methods: placement and tracking of landmarks, edge detection, and binary thresholding. /Scale: NFCS, PIPP and DAN	Own Base - Ordine Mauriziano Hospital (15 healthy full-term neonates between 48 and 72 hours of life).	Eye squeeze (between mid-eyebrow and mid-lower eyelid), cheek raise (between eye medial corner and nose corner), brow bulging (between eyebrows medial border).	The overall result is not reported, but some operators' evaluations were particularly inconsistent regarding some parameters like face furrowing. For these parameters, the scores had very low consistency (about 40%).
39	The commonly used face detection methods are introduced first, and then, the convolutional neural network in deep learning is analyzed and improved and then applied to the facial recognition of newborns. /Scale: Used in Hubei hospital	Own Base - Hubei hospital (40 newborns with the age of no more than 7 days).	Not applicable.	The accuracy of the improved algorithm is 0.889, higher by 0.036 in contrast to other models; the area under the curve (AUC) of success rate reaches 0.748, higher by 0.075 compared with other algorithms.
40	Compare five pre-trained face detection models, proposing two new NICUface models. /Scale: Unmentioned	CHEO (33 newborns), COPE (27 newborns), and NBHR (257 patients)	Not applicable.	The proposed NICUface models outperform previous state-of-the-art models for neonatal face detection and are robust to many identified complex NICU scenes.

Note: APN-db, Acute Pain in Neonates; AUC, Area Under the Curve; BPSN, Bernese Pain Scale for Neonates; CHEO, Children's Hospital of Eastern Ontario; COPE, Classification of Pain Expression; CRIES, C—Crying; R—Requires increased oxygen administration; I—Increased vital signs; E—Expression; S—Sleeplessness; DAN, Douleur Aigue Nouveau-Né; FLACC, Face, Legs, Activity, Cry, Consolability; Fuzzy k-NN, Fuzzy K-Nearest Neighbor; HOG, Histogram of Oriented Gradients; HOMO, Homomorphic Filter; IDE, Integrated Development Environment; LBP, Local Binary Pattern; NBHR, Newborn Baby Heart Rate; N-CNN, Neonatal - Convolutional Neural Networks; NFCS, Neonatal Facial Coding System; NIPS, Neonatal Infant Pain Scale; NPA, Neonatal Pain Assessment; NPAD, Neonatal Pain Assessment Dataset; N-PASS, Neonatal Pain and Sedation Scale; PIPP, Premature Infant Pain Profile; PIPP-R, Premature Infant Pain Profile-Revised; RMSE, Root Mean Square Error; UNIFESP, Universidade Federal de São Paulo; VAS, Visual Analogue Scale; VGG, Visual Geometry Group.

Table 4 Result of the literature search.

Reference	Each study's limitation	Reported perspectives
¹⁴	• The software is unable to detect points corresponding to the lower area of the face, such as chin movements, or specific points on the tongue; • The face with medical devices is not detected.	• It should be noted that the ideal tool would indicate when pain intensity deserves treatment or treatment adjustments; • The use of an electronic eye, less dependent on humans, could help to integrate pain assessment with pain treatment in the context of neonatal care.
²³	• Translational research needs to be done to assess the accuracy of the app for different neonates and clinical situations to assess if the performance of the neural network is less prone to subjective variables that modify pain assessment than the human performance; • The face with medical devices is not detected.	• Apply different explainable AI methods to better understand facial regions that might be relevant to pain assessment and use all the depicted face images to enlarge the face image samples and train our computational model. • Evaluate more recent CNN architectures. • Perform hands-on testing of the mobile application.
³⁰	• Not applicable.	• Not applicable.
³⁰ cited ²⁰	• The current approach was evaluated on a relatively small number of infants; • The face with medical devices is not detected.	• Explore the use of CNNs to develop a highly accurate pain assessment application; • Improve the effectiveness of pain intervention while mitigating the short- and long-term outcomes of pain exposure early in life; • Realize a multimodal approach to pain assessment that allows pain to be assessed during circumstances when all pain responses are not available, clinical condition, activity level, and sedation; • Integrate contextual information, such as medication type/dose, to obtain a context-sensitive pain assessment. • Collect a large multimodal dataset during hospitalization in the NICU; • Investigate the possibility of using neonate sounds as soft biometrics.
³⁰ cited ²⁶	• High False Negative Rate; • The current approach was evaluated on a relatively small number of infants; • The current work does not provide a comparison between the assessment of the proposed automatic system and human; • The face with medical devices is not detected.	• Investigate several directions for minimizing False Negative Rate; • Employ or implement advanced noise reduction methods; • Enlarge the training data using traditional augmentation methods and Generative adversarial networks; • Follow another approach, assessing the level or intensity of the detected pain class; • Evaluate the approach on a larger dataset of infants recorded during both procedural and postoperative pain; • Investigate the association between neonatal pain and the brain's hemodynamic activities using Near-infrared Spectroscopy; • Explore the association between neonatal pain and changes in skin color as well as the association between pain and eye movement/pupil dilation; • Test how well the automatic system performs as compared to human judgments; • Adding points to the total score of infants, based on age, pain history, or other factors, to compensate for their limited ability to behaviorally or physiologically communicate pain.

Table 4 (Continued)

Reference	Each study's limitation	Reported perspectives
31	• In terms of PCC, the performance is relatively low; • The face with medical devices is not detected.	• The goal is pain intensity estimation; • Future work will focus on incorporating other pain modalities, particularly body movements which are part of the NFLAPS and collecting additional data.
32	• Does not report whether the system will be able to evaluate faces with medical devices.	• Providing the relatives (parents) with a tool that allows remote supervising of their newborn's wellness.
33	• The face with medical devices is not detected.	• Evaluate the intensity of the pain, classifying it as mild, moderate, or severe pain; • Future applications may also include patient populations incapable of expressing pain (children with disability, adults with dementia, or mechanically ventilated patients).
34	• It is not possible to accurately assess pain scores or pain grades with AI technology; • The automated NPA system currently requires an additional nurse to record the video; • Does not report whether the system will be able to evaluate faces with medical devices.	• Automate the entire process by recording video with bedside cameras in the future. • The AI technology embedded in the electronic-medical-record system in the future will realize intervention for pain in real-time by medical staff. • The downstream health education system can further perform pain-knowledge education for newborns' family members to realize the traceability, standardization, and intelligence of the whole process of pain management.
35	• The metrics to validate the presence of artifacts need to be further tested. Mutual information was not able to validate, Pearson's correlation coefficient validated a part of the cases, and only the root means square error showed promise.	• Test more metrics and techniques to detect the presence of artifacts in each facial region.
36	• The system should be more illumination independent and should be able to handle partial component occlusions. • The face with medical devices is not detected.	• Perform an analysis possible to be applied in a real hospital situation.
37	• The face with medical devices is not detected.	• Implement the system in clinical practice, since the classification result shows that the proposed technique could be employed as a valuable tool for classifying the newborn between pain and normal with Fuzzy k-NN Classifier.
38	• The limited size of the dataset requires extension and further experimentation; • Rapid head movements are a problem for landmark tracking; • A thorough exploration of the parameters and alternatives of the KLT algorithm seems necessary to ensure more stability to the system; • The face with medical devices is not detected.	• Conduct additional video acquisition campaigns that will lead to a substantial extension of the original dataset; • Increase the number of expert operators performing manual pain assessment. • To better analyze the consistency and repeatability of the process, multiple scoring sessions on the videos should be performed by the same operators at different periods; • Adopt landmark selection algorithms best suited for child's faces; • Video processing can be effectively complemented by the analysis of audio information.

Table 4 (Continued)

Reference	Each study's limitation	Reported perspectives
39	• The model cannot be updated online; • The amount of data in the newborn image database is not very large, and it must amplify the image on the data set during the experiment; • It needs more experiments with faces covered by devices.	• In the follow-up research, the algorithm can be further optimized based on the model algorithm built by this research firstly to improve its real-time performance and recognize the emotions of newborns in real-time. • The algorithm can combine with other auxiliary information such as crying and body movements for multi-modal classification and recognition when the facial expressions of newborns are collected to build a standard newborn image database with a large volume of data.
40	• Not all methods were able to robustly identify patients' faces in complex scenes involving phototherapy lighting, ventilation support, view in dorsal decubitus, and prone position; • One of the proposed settings works better on smaller faces, and the other works better in a brighter environment.	• Perform an enhancement to improve the accuracy of NICU-face; Complement the networks using an ensemble network and combining the strengths of the models; • Use these models for the implementation of other neonatal monitoring applications (e.g., in-home monitoring or intelligent monitoring applications from smartphones)

Note: CNN, convolutional neural networks; Fuzzy k-NN, Fuzzy K-Nearest Neighbor; KLT, Kanade–Lucas–Tomasi; NFLAPS, Neonatal Face and Limb Acute Pain; NICU, Neonatal Intensive Care Unit; NPA, Neonatal Pain Assessment; PCC, Pearson Correlation Coefficient.

These preliminary results revealed that using behavioral indicators of pain along with physiological indicators could better assess neonatal pain.¹⁵

In 2018, researchers created a computational framework for pattern detection, interpretation, and classification of frontal face images for automatic pain identification in neonates. Classification of pain faces by the computational framework versus classification by healthcare professionals using the pain scale "Neonatal Facial Coding System" reached 72.8% accuracy. The authors reported that some disagreements between the assessment methods could stem from unstudied confounding factors, such as the classification of faces related to stress or newborn discomfort.¹⁶

In the same year, another group of researchers presented a dynamic method related to the duration of facial activity, by combining temporal and spatial representations of the face.¹⁷ In this study, the authors used facial configuration descriptors, head pose descriptors, numerical gradient descriptors, and temporal texture descriptors to describe facial changes over time. The dynamic facial representation and the multi-feature combination scheme were successfully applied for infant pain assessment. The authors concluded that the profile-based infant pain assessment is also feasible because its performance was almost as good as using the whole face. In addition, the authors noted that gestational age was one of the most influencing factors for infant pain assessment, highlighting the importance of designing specific models depending on gestational age.¹⁷

Other researchers have implemented a computational framework using triangular meshes to generate a spatially normalized atlas of high resolution, potentially useful for the automatic evaluation of neonatal pain.¹⁹ These atlases are essential to describe characteristic and detailed facial

patterns, preventing image effects or signals (which are not relevant and which portray undesirable particularities, inherent to the imperfect data acquisition process) from being erroneously propagated as discriminative variations.

Also in 2019, researchers created a network for neonatal pain classification, called Neonatal - Convolutional Neural Network (N-CNN), designed to analyze neonates' facial expressions. The proposed network achieved encouraging results that suggested that automated neonatal pain recognition may be a viable and efficient alternative for pain assessment.²⁰ This was the first CNN built specifically for neonatal pain assessment, which did not use transfer learning as a methodology.

In addition, another group of studies developed an automated neonatal discomfort detection system based on video monitoring, divided into two stages: (1) face detection and face normalization; (2) feature extraction and facial expression classification to discriminate infant status into comfort or discomfort. The experimental results showed an accuracy of 87% to 97%. However, even though the results were promising for use in clinical practice, the authors reported the need for new studies with more newborn data to evaluate and validate the system.²¹

A major technological advance occurred in 2020 after the development of a new video dataset for automatic neonatal pain detection called iCOPEvid (infant Classification of Pain Expressions videos). The creators of this dataset also presented a system to classify the iCOPEvid segments into two categories: pain and non-pain. Compared to other human classification systems, the results were superior; however, the addition of CNN to further improve the results was not successful. Therefore, the authors reported the need for further studies using CNN.¹⁸

In 2020, a new study proposed an application of multivariate statistical analysis, in the context of images of newborns with and without pain, to explore, quantify, and determine behavioral measures that would help in the creation of generalist pain classification models, both by automated systems and by health professionals. The authors reported that using behavioral measures it was possible to classify the intensity of pain expression and identify the main facial regions involved in this process (frowning the forehead, squeezing the eyes, deepening the nasolabial groove, and horizontally opening the mouth made the model similar to a face with pain, and features such as mouth closure, eye-opening, and forehead relaxation made the model similar to a face without pain). The developed framework showed that it is possible to statistically classify the expression of pain and non-pain through facial images and highlight discriminant facial regions for the pain phenomenon.¹⁹

In 2021, two studies were conducted using deep neural networks. One compared the use of the N-CNN and an adapted ResNet50 neural network architecture to find the model best suited to the neonatal face recognition task. The results showed that the modified ResNet50 model was the best one, with an accuracy of 87.5% for the COPE image bank.²² The other study used neural networks for newborn face detection and pain classification in the context of mobile applications. Additionally, this was the first study to apply explainable Artificial Intelligence (AI) techniques in neonatal pain classification.²³

Then, new research reviewed the practices and challenges for pain assessment and management in the NICU using AI. The researchers reported that AI-based frameworks can use single or multiple combinations of continuous objective variables, that is, facial and body movements, cry frequencies, and physiological data (vital signs) to make high-confidence predictions about the time-to-pain onset following postsurgical sedation. The authors reported that emerging AI-based strategies have the potential to minimize or avoid damage to the newborn's body and psyche from postsurgical pain and opioid withdrawal.²⁴

Another study group has created an AI System, called "PainChek Infant" for automatic recognition and analysis of the face of infants aged 0 to 12 months, allowing the detection of six facial action units indicative of the presence of pain. PainChek Infant pain scores showed a good correlation with "Neonatal Facial Coding System-R" and the "Observer-administered Visual Analogue Scale" scores ($r = 0.82-0.88$; $p < 0.0001$). PainChek Infant also showed good to excellent interrater reliability ($ICC = 0.81-0.97$, $p < 0.001$) and high levels of internal consistency ($\alpha = 0.82-0.97$).²⁵

In 2022, a pain assessment system was created using facial expressions, crying, body movement, and vital sign changes. The proposed automatic system generated a standardized pain assessment comparable to those obtained by conventional nurse-derived pain scores. According to the authors, the system achieved 95.56% accuracy. The results showed that the automatic assessment of neonatal pain is a viable and more efficient alternative than the manual assessment.²⁶

Additionally, in 2023, a systematic review study discussed the models, methods, and data types used to lay the foundations for an automated pain assessment system based on deep learning. In total, one hundred and ten pain assessment works based on unimodal and multimodal approaches

were identified for different age groups, including neonates. According to the authors, artificial intelligence solutions in general, and deep neural networks in particular, are models that perform complex functions, but lack transparency, which becomes the main reason for criticism. Also, this review demonstrated the importance of multimodal approaches for automatic pain estimation, especially in clinical settings, and highlights that the limited number of studies exploring the phenomenon of pain beyond extraordinary situations, or considering different contexts, maybe one of the limitations of current approaches regarding their applicability in real-life settings and circumstances.⁴³

All studies reported significant limitations that preclude the use of their methods in the clinical NICU practice, such as (1) the inability to detect points corresponding to the lower facial area, chin movements, specific tongue points, and rapid head movements; (2) a small number of neonates for evaluation and testing of the algorithm; (3) inability to robustly identify patients' faces in complex scenes involving lighting and ventilation support. Given these limitations, there is an emerging need for evaluating and validating each neonatal pain assessment automation method proposed to date.

A limited number of databases of neonatal facial images

The reason for the small number of published studies around automated neonatal pain analysis and assessment using Computer Vision and Machine Learning technologies may be related to the limited number of neonatal image datasets available for research.¹⁷

Currently, there are few datasets for facial expression analysis of pain in newborns. The publicly available databases are COPE;⁴¹ Acute Pain in Neonates database (APN-db);³¹ Facial Expression of Neonatal Pain (FENP);⁴⁴ freely available data from YouTube, which was used from the year 2014 for a systematic review study;⁴⁵ USF-MNPAD-I (University of South Florida Multimodal Neonatal Pain Assessment Dataset);⁴⁶ and Newborn Baby Heart Rate Estimation Database (NBHR)⁴⁷ which provides facial images of newborns, but is primarily aimed at monitoring physiological signs.

All these databases are being widely used in the academic scientific environment; however, they have some limitations, such as the small number of images; images of only one ethnic group; low confidence (no explanation about approval in ethics committees); images of a specific clinical population, mainly term newborns, not allowing studies with preterm and critically ill newborns.

These databases only have images of newborns with the face free and do not have images of newborns with devices attached to the face, except for the USF-MNPAD-I database.⁴⁶ This scenario of a scarcity of databases with images of critically ill newborns hampers the development of new methods to automate pain assessment in this very specific population.

Perspectives

This article attempted to report the difficulties faced so far in the creation of an automatic method for pain assessment in the neonatal context. Based on the several analyzed frameworks, it is evident that there are gaps in the

development of practical applications that are sensible, specific, with good accuracy, and can be used at the bedside.

For research to advance in this area, a larger number of neonatal facial images are needed to test and validate algorithms. The authors believe that a convenient way to overcome this practical issue would be the creation of synthetic databases, which might contemplate not only the increased number of facial images but also different races, sex, types of patients, and types of devices used, aiming at a better generalization of the algorithms.

Another limitation in the studies is the difficulty of detecting pain in partially covered faces. As previously stated, the devices attached to the face of the newborn hinder the visualization of all points and facial regions that are indispensable for the automatic evaluation of pain.

This problem only happens because the computational methods developed so far assume that all facial regions need to be detected, evaluated, and scored to make the pain diagnosis possible. Trying to guess what the image of the facial region behind the medical device looks like may not be the best alternative to solve this problem.

One possibility would be to identify pain only by analyzing the free facial regions. The development of a system of evaluation of the segmented parts of the face would make possible the evaluation and classification of pain weighted only by the free facial regions, not requiring the identification and classification of all facial points, as is done holistically nowadays.

The creation of a classifier by facial region would allow the identification of which regions are more discriminating for the diagnosis of neonatal pain. Consequently, it would be possible to give scores with different weights for each visible facial region and maximize the process of pain assessment of newborns who remain with part of the face occluded.

In addition, new methods of facial assessment need to be tested during the natural movement of the neonate and with different light intensities. It would be important to test how the computer method for automatic pain assessment works together with other assessment methods such as manual assessment using facial pain scales, body movement assessment, brain activity, sweating, skin color, pupil dilation, vital signs, and crying.

It is worth mentioning that, for decision-making, neither clinical practice (based on pain scales) nor computer models alone would be sufficient to reach a more accurate decision process. This is because, without interpreting the information used for decision-making by humans and machines, one cannot affirm that the assessment was made with precision. Therefore, studies that seek to understand this information, extracted from both human and machine eyes, can help to create models that combine these two types of learning. Examples of such studies would be those of Silva et al.⁴⁸ Barros et al.⁴⁹ and Soares et al.⁵⁰ that used gaze tracking of observers during newborn pain assessment; and research using eXplainable Artificial Intelligence (XAI) models, such as those of Carlini et al.²³ and Coutin et al.⁵¹

With technological advances, it will be possible to create a method capable of identifying the presence and the intensity of neonatal pain, differentiating pain from discomfort and acute pain from chronic pain. Thus, providing the appropriate neonatal care and treatment for each patient, according to the gestational age and within the complexity that involves the NICU environment.

Funding source

This work received financial support from the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior CAPES related to the doctoral scholarship of the student Tatiany Marcondes Heiderich (Brazil/Process: n. 142499/2020-0) and the Centro Universitário da FEI - Fundação Educacional Inaciana.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgments

To CAPES and Centro Universitário da FEI - Fundação Educacional Inaciana for the scholarships.

References

1. Grunau RVE, Craig KD. Pain expression in neonates: facial action and cry. *Pain*. 1987;28:395–410.
2. Guinsburg R. Assessing and treating pain in the newborn. *J Pediatr*. 1999;75:149–60.
3. Stevens BR, Johnston CR, Petryshen PR, Taddio AB. Premature infant pain profile: development and initial validation. *Clin J Pain*. 1996;12:13–22.
4. Stevens BJ, Gibbins S, Yamada J, Dionne K, Lee G, Johnston C, et al. The Premature Infant Pain Profile-Revised (PIPP-R). *Clin J Pain*. 2014;30:238–43.
5. Hummel P, Puchalski M, Creech S, Weiss MG. N-PASS: neonatal pain, agitation, and sedation scale - reliability and validity. In: Presented at: Pediatric Academic Societies' annual meeting. Seattle, WA; 2003.
6. Debillon T, Zupan V, Ravault N, Magny JF, Dehan M. Development and initial validation of the EDIN scale, a new tool for assessing prolonged pain in preterm infants. *Arch Dis Child Fetal Neonatal Ed*. 2001;85:36F–41.
7. Krechel SW, Mc JB. CRIES: a new neonatal postoperative pain measurement score. Initial testing of validity and reliability. *Paediatr Anaesth*. 1995;5:53–61.
8. van Dijk M, Roofthoof DW, Anand KJ, Guldemon F, de Graaf J, Simons S, et al. Taking up the challenge of measuring prolonged pain in (premature) neonates. *Clin J Pain*. 2009;25:607–16.
9. Hand IL, Noble L, Geiss D, Wozniak L, Hall C. COVERS Neonatal pain scale: development and validation. *Int J Pediatr*. 2010;2010:496719.
10. Hudson-Barr D, Capper-Michel B, Lambert S, Mizell Palermo T, Morbeto K, Lombardo S. Validation of the Pain Assessment in Neonates (PAIN) Scale with the Neonatal Infant Pain Scale (NIPS). *Neonatal Netw*. 2002;21:15–21.
11. Lawrence J, Alcock D, McGrath P, Kay J, MacMurray SB, Dulberg C. The development of a tool to assess neonatal pain. *Neonatal Netw*. 1993;12:59–66.
12. Zeng Z. Assessment of neonatal pain: uni- and multidimensional evaluation scales. *Front Nurs*. 2022;9:247–54.
13. Serpa AB, Guinsburg R, Balda Rde C, dos Santos AM, Areco KC, Peres CA. Multidimensional pain assessment of preterm newborns at the 1st, 3rd and 7th days of life. *Sao Paulo Med J*. 2007;125:29–33.
14. Heiderich TM, Leslie AT, Guinsburg R. Neonatal procedural pain can be assessed by computer software that has good sensitivity

- and specificity to detect facial movements. *Acta Paediatr.* 2015;104:e63–9.
15. Zamzmi G, Pai CY, Goldgof D, Kasturi R, Ashmeade T, Sun Y. An approach for automated multimodal analysis of infants' pain. In: 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE; 2016:4148–53. [Cited 2023 Mar 8]. Available from: <http://ieeexplore.ieee.org/document/7900284/>.
 16. Teruel GF, Heiderich TM, Guinsburg R, Thomaz CE. Analysis and recognition of pain in 2d face images of full term and healthy newborns. *Anais do XV Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2018). Sociedade Brasileira de Computação - SBC; 2018. p. 228–39.* [Cited 2023 Mar 8]. Available from: <http://portaldeconteudo.sbc.org.br/index.php/eniac/article/view/4419>.
 17. Zhi R, Zamzmi G, Goldgof D, Ashmeade T, Sun Y. Automatic infants' pain assessment by dynamic facial representation: effects of profile view, gestational age, gender, and race. *J Clin Med.* 2018;7:173.
 18. Brahnam S, Nanni L, McMurtrey S, Lumini A, Brattin R, Slack M, et al. Neonatal pain detection in videos using the iCOPEvid dataset and an ensemble of descriptors extracted from Gaussian of Local Descriptors. *Appl Comput Inform.* 2020: 1–22.
 19. Orona PA, Fabbro DA, Heiderich TM, Barros MC, RdeC Balda, Guinsburg R, et al. Atlas of neonatal face images using triangular Meshes. *Anais do XV Workshop de Visão Computacional (WVC 2019). Sociedade Brasileira de Computação - SBC; 2019. p. 19–24.* [Cited 2023 Mar 8]. Available from: <https://sol.sbc.org.br/index.php/wvc/article/view/7622>.
 20. Zamzmi G, Paul R, MdS Salekin, Goldgof D, Kasturi R, Ho T, et al. Convolutional neural networks for neonatal pain assessment. *IEEE Trans Biom Behav Identity Sci.* 2019;1:192–200.
 21. Sun Y, Shan C, Tan T, Long X, Pourtaherian A, Zinger S, et al. Video-based discomfort detection for infants. *Mach Vis Appl.* 2019;30:933–44.
 22. Buzuti L, Heideirich T, Barros M, Guinsburg R, Thomaz C. Neonatal pain assessment from facial expression using deep neural networks. *Anais do XVI Workshop de Visão Computacional (WVC 2020). Sociedade Brasileira de Computação - SBC; 2020. p. 87–92.* [Cited 2023 Mar 8]. Available from: <https://sol.sbc.org.br/index.php/wvc/article/view/13486>.
 23. Carlini LP, Ferreira LA, Coutin GAS, Varoto VV, Heiderich TM, Balda RCX, et al. A convolutional neural network-based mobile application to bedside neonatal pain assessment. In: 2021 34th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), IEEE; 2021:394–401. [Cited 2023 Mar 8]. Available from: <https://ieeexplore.ieee.org/document/9643144/>.
 24. Salekin MS, Mouton PR, Zamzmi G, Patel R, Goldgof D, Kneusel M, et al. Future roles of artificial intelligence in early pain management of newborns. *Paediatr Neonatal Pain.* 2021;3:134–45.
 25. Hoti K, Chivers PT, Hughes JD. Assessing procedural pain in infants: a feasibility study evaluating a point-of-care mobile solution based on automated facial analysis. *Lancet Digit Health.* 2021;3:e623–34.
 26. Zamzmi G, Pai CY, Goldgof D, Kasturi R, Ashmeade T, Sun Y. A comprehensive and context-sensitive neonatal pain assessment using computer vision. *IEEE Trans Affect Comput.* 2022;13:28–45.
 27. Graziosi MES, Liebano RE, Nahas FX, Pesquisa em Bases de Dados - Módulo Científico, In: Especialização em Saúde da Família UNA-SUS, 25–33, [cited 2022 Aug 14]. Available from: https://www.unasus.unifesp.br/biblioteca_virtual/esf/1/modulo_cientifico/Unidade_13.pdf
 28. EESC - USP. Revistas Científicas na Área da Engenharia . [cited 2022 Aug 30]. Available from: https://eesc.usp.br/biblioteca/post.php?guid=95&catid=fonte_eletronica.
 29. Biblioteca Virtual em Saúde - BVS. DeCS/MeSH - Descritores em Ciências da Saúde. [cited 2022 Aug 30]. Available from: <https://decs.bvsalud.org/>.
 30. Grifantini K. Detecting faces, saving lives. *IEEE Pulse.* 2020;11:2–7. [cited 2022 Aug 14]. Available from: <https://ieeexplore.ieee.org/document/9089065/>. [cited 2022 Aug 14]. Available from: .
 31. Egede J, Valstar M, Torres MT, Sharkey D. Automatic neonatal pain estimation: an acute pain in neonates database. In: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII), IEEE; 2019:1–7. [Cited 2023 Mar 8]. Available from: <https://ieeexplore.ieee.org/document/8925480/>.
 32. Martinez-Balleste A, Casanovas-Marsal JO, Solanas A, Casino F, Garcia-Martinez M. An autonomous system to assess, display and communicate the pain level in newborns. In: 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA), IEEE; 2014:1–5. [Cited 2023 Mar 8]. Available from: <http://ieeexplore.ieee.org/document/6860144/>.
 33. Roué JM, Morag I, Haddad WM, Gholami B, Anand KJS. Using sensor-fusion and machine-learning algorithms to assess acute pain in non-verbal infants: a study protocol. *BMJ Open.* 2021;11:e039292.
 34. Cheng X, Zhu H, Mei L, Luo F, Chen X, Zhao Y, et al. Artificial intelligence based pain assessment technology in clinical application of real-world neonatal blood sampling. *Diagnostics.* 2022;12:1831.
 35. Domingues PH, da Silva RM, Orra IJ, Cruz ME, Heiderich TM, Thomaz CE. Neonatal face mosaic: an areas-of-interest segmentation method based on 2D face images. *Anais do XVII Workshop de Visão Computacional (WVC 2021). Sociedade Brasileira de Computação - SBC; 2021. p. 201–5.* [Cited 2023 Mar 8]. Available from: <https://sol.sbc.org.br/index.php/wvc/article/view/18914>.
 36. Han J, Hazelhoff L, de With PHN. Neonatal monitoring based on facial expression analysis. *Neonatal Monitoring Technologies. IGI Global; 2012. p. 303–23.* [Cited 2023 Mar 8]. Available from: <http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-0975-4.ch014>.
 37. Mansor MN, Junoh AK, Ahmed A, Kamarudin H, Idris A. Infant pain detection with homomorphic filter and fuzzy k-NN classifier. *Appl Mech Mater.* 2014;643:183–9. [Cited 2023 Mar 8]. Available from: <https://www.scientific.net/AMM.643.183>.
 38. Parodi E, Melis D, Boulard L, Gavelli M, Baccaglini E. Automated newborn pain assessment framework using computer vision techniques. In: In: Proceedings of the International Conference on Bioinformatics Research and Applications 2017 - ICBRA 2017, New York, New York, USA: ACM Press; 2017:31–6. [Cited 2023 Mar 8]. Available from: <http://dl.acm.org/citation.cfm?doid=3175587.3175590>.
 39. Wang Y, Huang L, Yee AL. Full-convolution Siamese network algorithm under deep learning used in tracking of facial video image in newborns. *J Supercomput.* 2022;78:14343–61.
 40. Dosso YS, Kyrollos D, Greenwood KJ, Harrold J, Green JR. NICU-face: Robust neonatal face detection in complex NICU scenes. *IEEE Access.* 2022;10:62893–909. [Cited 2023 Mar 8]. Available from: <https://ieeexplore.ieee.org/document/9791241/>.
 41. Brahnam S, Chuang CF, Shih FY, Slack MR. Machine recognition and representation of neonatal facial displays of acute pain. *Artif Intell Med.* 2006;36:211–22.
 42. Schiavenato M, Byers JF, Scovanner P, McMahon JM, Xia Y, Lu N, et al. Neonatal pain facial expression: Evaluating the primal face of pain. *Pain.* 2008;138:460–71.
 43. Gkikas S, Tsiknakis M. Automatic assessment of pain based on deep learning methods: a systematic review. *Comput Methods Programs Biomed.* 2023;231:107365.
 44. Yan J, Lu G, Li X, Zheng W, Huang C, Cui Z, et al. FENP: a database of neonatal facial expression for pain analysis. *IEEE Trans Affect Comput.* 2023;14:245–54.

45. Harrison D, Sampson M, Reszel J, Abdulla K, Barrowman N, Cumber J, et al. Too many crying babies: a systematic review of pain management practices during immunizations on YouTube. *BMC Pediatr.* 2014;14:134.
46. Salekin MS, Zamzmi G, Hausmann J, Goldgof D, Kasturi R, Kneusel M, et al. Multimodal neonatal procedural and postoperative pain assessment dataset. *Data Brief.* 2021;35:106796.
47. Huang B, Chen W, Lin CL, Juang CF, Xing Y, Wang Y, et al. A neonatal dataset and benchmark for non-contact neonatal heart rate monitoring based on spatio-temporal neural networks. *Eng Appl Artif Intell.* 2021;106:104447. [Cited 2023 Mar 8]. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0952197621002955>.
48. Silva GV, Barros MC, Soares JD, Carlini LP, Heiderich TM, Orsi RN, et al. What facial features does the pediatrician look to decide that a newborn is feeling pain? *Am J Perinatol.* 2023;40:851–7.
49. Barros MC, Thomaz CE, da Silva GV, do Carmo AS, Carlini LP, Heiderich TM, et al. Identification of pain in neonates: the adults' visual perception of neonatal facial features. *J Perinatol.* 2021;41:2304–8.
50. Soares JD, Barros MC, da Silva GV, Carlini LP, Heiderich TM, Orsi RN, et al. Looking at neonatal facial features of pain: do health and non-health professionals differ? *J Pediatr (Rio J).* 2022;98:406–12.
51. Coutrin GAS, Carlini LP, Ferreira LA, Heiderich TM, Balda RCX, Barros MCM, et al. Convolutional neural networks for newborn pain assessment using face images: A quantitative and qualitative comparison. In: *3rd International Conference on Medical Imaging and Computer-Aided Diagnosis - MICAD 2022*, Leicester, UK: Springer LNEE; 2022.