

7-2018

## Automatic Infants' Pain Assessment by Dynamic Facial Representation: Effects of Profile View, Gestational Age, Gender, and Race

Ruicong Zhi

*University of Science and Technology Beijing*

Ghada Z. D. Zamzmi

*University of South Florida, ghadh@mail.usf.edu*

Dmitry Goldgof

*University of South Florida, goldgof@mail.usf.edu*

Terri Ashmeade

*University of South Florida, tashmead@health.usf.edu*

Yu Sun

*University of South Florida, yusun@cse.usf.edu*

Follow this and additional works at: [https://scholarcommons.usf.edu/esb\\_facpub](https://scholarcommons.usf.edu/esb_facpub)

---

### Scholar Commons Citation

Zhi, Ruicong; Zamzmi, Ghada Z. D.; Goldgof, Dmitry; Ashmeade, Terri; and Sun, Yu, "Automatic Infants' Pain Assessment by Dynamic Facial Representation: Effects of Profile View, Gestational Age, Gender, and Race" (2018). *Computer Science and Engineering Faculty Publications*. 115.

[https://scholarcommons.usf.edu/esb\\_facpub/115](https://scholarcommons.usf.edu/esb_facpub/115)

This Article is brought to you for free and open access by the Computer Science and Engineering at Scholar Commons. It has been accepted for inclusion in Computer Science and Engineering Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact [scholarcommons@usf.edu](mailto:scholarcommons@usf.edu).



Article

# Automatic Infants' Pain Assessment by Dynamic Facial Representation: Effects of Profile View, Gestational Age, Gender, and Race

Ruicong Zhi <sup>1,2,\*</sup>, Ghada Zamzmi Dmitry Goldgof <sup>3</sup>, Terri Ashmeade <sup>4</sup> and Yu Sun <sup>3</sup>

<sup>1</sup> School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China

<sup>2</sup> Beijing Key Laboratory of Knowledge Engineering for Materials Science, Beijing 100083, China

<sup>3</sup> Department of Computer Science and Engineering, University of South Florida, Tampa, FL 33620, USA; ghadh@mail.usf.edu (G.Z.); goldgof@mail.usf.edu (D.G.); yusun@mail.usf.edu (Y.S.)

<sup>4</sup> College of Medicine Pediatrics, University of South Florida, Tampa, FL 33620, USA; tashmead@health.usf.edu

\* Correspondence: zhirc\_research@126.com; Tel.: +86-138-1048-8859

Received: 28 May 2018; Accepted: 3 July 2018; Published: 11 July 2018



**Abstract:** Infants' early exposure to painful procedures can have negative short and long-term effects on cognitive, neurological, and brain development. However, infants cannot express their subjective pain experience, as they do not communicate in any language. Facial expression is the most specific pain indicator, which has been effectively employed for automatic pain recognition. In this paper, dynamic pain facial expression representation and fusion scheme for automatic pain assessment in infants is proposed by combining temporal appearance facial features and temporal geometric facial features. We investigate the effects of various factors that influence pain reactivity in infants, such as individual variables of gestational age, gender, and race. Different automatic infant pain assessment models are constructed, depending on influence factors as well as facial profile view, which affect the model ability of pain recognition. It can be concluded that the profile-based infant pain assessment is feasible, as its performance is almost as good as that of the whole face. Moreover, gestational age is the most influencing factor for pain assessment, and it is necessary to construct specific models depending on it. This is mainly because of a lack of behavioral communication ability in infants with low gestational age, due to limited neurological development. To our best knowledge, this is the first study investigating infants' pain recognition, highlighting profile facial views and various individual variables.

**Keywords:** infants' pain assessment; dynamic facial representation; profile view; gestational age; gender; race

## 1. Introduction

Healthcare for infants in a Neonatal Intensive Care Unit (NICU) is critical for survival; however, the hospitalization stage affects their neurodevelopment and future growth. Invasive medical interventions are often required in clinical treatment, and as a result, infants suffer from repeated procedural pain as part of their general care in the NICU. The early exposure to painful procedures can have negative short and long-term effects on cognitive, neurological, and brain development [1]. Therefore, it is important to judge and measure pain procedures for infants, especially preterm infants under special medical care.

Accurate pain assessment is comprehensive and multidimensional, and the commonly used pain measurements include contextual, behavioral, and physiological tools [2]. Pain is subjective experience,

which may easily be affected by physical, cognitive, emotional state, and pain history [3,4]. Infants exhibit pain-related facial activities [5], body movements [6], physiological responses [7], and cortical activities [8] during painful procedures. A self-reported pain assessment method is not applicable for infants with communicative impairment. Multidimensional indicator-based scales are commonly used for infants' pain assessment, such as Neonatal Infant Pain Scale (NIPS) [9], Face, Legs, Activity, Crying, and Consolability (FLACC) [10], Neonatal Facial Coding System (NFCS) [11] for acute pain assessment, Neonatal Pain, Agitation, and Sedation Scale (N-PASS) [12], Neonatal Pain and Discomfort Scale (EDIN) [13], and Crying, Requires O<sub>2</sub>, Increased VS, Expression, and Sleepless (CRIES) [14] for chronic pain assessment.

The observational measures are based on several different behavioral indicators (e.g., facial expression, crying, body activity, and sleeping state) and physiological indicators (e.g., breathing pattern, heart rate, and oxygen saturation) related to pain. Since infants cannot communicate in words, observational measures are commonly preferred and considered as the gold-standard for pain assessment [15]. Although the indicator-based scales are easy to use in practice, the evaluation requires professional caregivers with plenty of training, and manual pain assessment is time-consuming and laboring for long-term continuous pain monitoring, as the caregiver has to assess infants' pain at short intervals. Moreover, such measures are easily disrupted by the observer's bias, and various influence factors such as clinical experience, underestimation of pain [16], background, and culture [17,18]. With the rapid development of machine learning and artificial intelligence, an automatic infant pain assessment system is desired for objective and accurate pain assessment.

## 2. Related Work and Contributions

There has been an increasing interest in understanding individual behavioral responses to pain based on facial expressions [19–23], body or head movements [24,25], and sound signals (crying) [26–28]. Pain-related behavior analysis is non-invasive and easily acquired by video recording technique. Evidence supports the fact that facial expression is the most specific indicator and is more salient and consistent than other behavioral indicators [29,30].

Facial expressions can provide insight into an individual's emotional state, and automatic facial expression analysis (AFEA) is a topic of broad research [31]. However, the contribution of pain expression is less extensive, especially for infant pain assessment. There are several researches on pain facial expression recognition in adults [19,20,23,32–34]. However, the methods designed for adult pain assessment may not show similar performance and may completely fail in infants for three main reasons: First, facial morphology and dynamics vary between infants and adults, as reported in [35]. Moreover, infant facial expressions include additional important units that are not present in the Facial Action Coding System. As such, NFACS (Neonatal Facial Coding System) is introduced as an extension of FACS [35,36]. Second, infants with different individual variables (such as gestational age) have different pain facial characteristics due to a less developed central nervous system [36]. Third, the preprocessing stage is more challenging in case of infants, since they are uncooperative subjects recorded in an unconstrained environment. In this paper, we focus on studies on infant pain assessment.

Brahnam et al. [37] utilized the holistic eigenfaces approach to recognize pain facial expressions in newborn babies, and compared the performances of distance-based classifier and Support Vector Machine (SVM) for pain detection. This work was extended by employing Sequential Floating Forward Selection for feature selection and Neutral Network Simultaneous Optimization Algorithm (NNSOA) for classification; an average classification rate of 90.2% was obtained [38]. Gholami et al. [39] presented Relevance Vector Machine (RVM) to assess infant pain and its intensity. The classification accuracy of RVM (85%) was found to be close to assessments by experts. Nanni et al. [40] used several histogram-based descriptors to detect infant pain facial expression, including Local Binary Pattern (LBP), Local Ternary Pattern (LTP), Elongated Ternary Pattern (ELTP), and Elongated Binary Pattern (ELBP). The highest accuracy was achieved by ELTP with Area under the Curve of Receiver Operating

Characteristic Curve (AUC) score of 0.93. The above researches were conducted using the COPE database, which is the only open access infant pain database, consisting of 204 static 2D images of 26 infants photographed under pain stimuli. Static images reveal facial expressions in certain photographs, while ignoring temporal information of pain.

A few studies have recently focused on dynamic pain facial expression analysis. Fotiadou et al. [41] applied the Active Appearance Model (AAM) to extract facial features and global motion for each video frame. SVM classifier was utilized for pain detection in 15 videos of eight infants, and the AUC achieved 0.98. Zamzmi et al. [42,43] extracted pain-relevant facial features from video sequences by estimating the optical strain magnitudes corresponding to pain facial expressions. SVM and KNN (K-nearest neighbor) classifiers were employed for pain detection, and an overall accuracy of 96% was obtained.

Most AFEA systems focus on facial expression analysis in near-frontal-view of facial recordings, and very few studies investigate the effect of the profile view of a facial image [44,45]. According to clinical observations, head movements occur commonly during pain experiences. Head shaking results in multi-view faces and may lead to failure of face detection and pain recognition. Facial expression recognition for a profile-view is challenging, as a lot of the facial representation information is lost. There is no research available that investigates pain facial expression recognition performances on profile view.

According to clinical research, infants with low gestational age have less-developed central nervous systems, and show limited ability to behaviorally communicate pain in comparison to full-term or post-term infants [46]. Derbyshire [47] revealed high pain sensitivity in female adults than in males with different situational cases, while pain responses for early age infants were not found to be extremely affected by sex differences. Due to the complexity of clinical contexture, infant pain facial expression analysis is more challenging, and contextual and individual factors are worth considering (i.e., age, gender, and race).

There is growing evidence in psychological research of the fact that temporal dynamics of facial behavior (e.g., the duration of facial activity) is a critical factor in the interpretation of observed behavior [45]. In this paper, we propose a dynamic pain facial expression representation and fusion scheme for automatic infant pain assessment, by combining temporal appearance facial features and temporal geometric facial features. Different automatic pain assessment models are constructed to gain a better understanding of the various factors that influence pain reactivity in infants, including gestational age, gender, and race. Moreover, a pain assessment model based on facial profile view is also investigated. The effectiveness of a specific model constructed according to individual variables is analyzed and compared with a general model. To the best of our knowledge, this is the first study investigating infant pain recognition depending on multiple facial views and various individual variables.

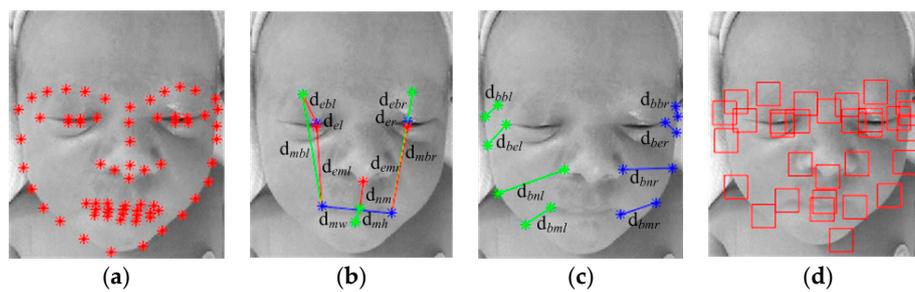
### 3. Methodology

Temporal information for facial representation is crucial for emotional analysis, especially for subtle facial expressions. Static images are photographed at discrete points in time, and they only include limited facial activity information where facial changes over time are missing; for example, eye blinking, which is important for pain recognition, cannot be distinguished by static features. In this section, we describe our dynamic pain facial activity representation method, including frame-level facial configuration/texture parameters representation, and sequence-based temporal descriptors extraction. Hybrid pain facial activity representation, including facial geometric features and texture features, is studied for pain facial expression recognition. The temporal pain-related facial activity representation is the same as in our preliminary work [48]. The dimensionality of hybrid facial features is reduced by manifold learning-based Supervised Locality Preserving Projections (SLPP) method, and SVM classifier is utilized for pain recognition. Finally, the classifier outputs of multiple facial activity features are combined by decision fusion.

### 3.1. Frame-Level Hybrid Facial Representation

The facial configuration parameters are designed by a set of geometric distance parameters, which intuitively depict facial deformation during pain experiences. On the other hand, facial appearance descriptors are derived from the gradient magnitude of a set of Region around Points (RAPs) centered on key facial fiducial points. Several facial activity representation parameters are extracted from each frame image and the details are described as follows.

- Facial landmark detection:** Both facial configuration parameters and facial texture parameters are calculated on the basis of facial landmarks. We apply the popular Active Appearance Model (AAM) to detect infants' facial landmarks (68 points) in image sequences to outline the main features of the eyebrows, eyes, nose, mouth, and face boundary (see Figure 1a). AAM combines a shape variation model with an appearance variation model in a shape-normalized frame, and is improved upon to be rapid, accurate, and robust [49].
- Facial deformation measurement:** Several pain-related geometric distance parameters are derived for each image frame from the facial fiducial points of the eyebrows, eyes, nose, and mouth, to capture deformations of facial components. Specific facial actions have been identified in Neonatal Facial Coding System (NFCS), which is a subsystem of Facial Action Coding System (FACS), and specially designed for pain assessment for newborns to 2 months of age. The validity and reliability of the sign judgment method has been evaluated in previous studies [3]. The facial actions for infant pain assessment include brow bulge, nasolabial furrow, eye squeeze, chin quiver, open lips, lip purse, horizontal mouth stretch, vertical mouth stretch, taut tongue, and tongue protrusion [36]. Thus, distance parameters are defined by the NFCS as the Euclidean distance between key facial fiducial points, including distance between the eyebrow and the eye ( $d_{ebl}$  and  $d_{ebr}$ ), distance between the upper eyelid and the lower eyelid ( $d_{el}$  and  $d_{er}$ ), distance between the eyebrow and the mouth ( $d_{mbl}$  and  $d_{mbr}$ ), distance between the eye and the mouth ( $d_{eml}$  and  $d_{emr}$ ), distance between the nose and the mouth ( $d_{nm}$ ), and the width ( $d_{mw}$ ) and height ( $d_{mh}$ ) of the mouth. Facial deformation parameters are shown in Figure 1b.
- Head pose movement measurement:** According to clinical observation, infants in pain usually shake their head left and right. Head movement is a useful predictor for infant pain assessment. Since there is no obvious relationship between pain occurrence/intensity and the head orientation angle, we conducted distance-based head pose parameters in a simple manner. The distance parameters include the distances between key facial components landmarks (eyebrows, eyes, nose, and mouth) and face boundary landmarks for both left and right sides of the face ( $d_{bbbl}$ ,  $d_{bbel}$ ,  $d_{bnl}$ ,  $d_{bnl}$ ,  $d_{bml}$ ,  $d_{bbr}$ ,  $d_{ber}$ ,  $d_{bnr}$ ,  $d_{bmr}$ ). The head pose parameters are illustrated in Figure 1c.
- Facial texture measurement:** Appearance-based facial expression features are extracted to reflect the magnitude and direction of skin surface displacement. Based on primary facial landmarks, patches with size of  $32 \times 32$  are defined around the landmarks. A domain knowledge-based texture description method is employed by calculating the mean magnitude for each local patch. These patches cover most of the facial regions, as illustrated in Figure 1d, where obvious wrinkles usually appear between the eyebrows (eyes), the corner of eyes, nasal root, and corner of the mouth. It is a simple and low-dimensional facial texture representation comparing to generic facial features, as only one numeric parameter is obtained for each RAP, and there are 31 texture parameters for each frame image.



**Figure 1.** Frame-level parameters of infant pain facial expression. (a) Facial landmarks; (b) facial configuration parameters; (c) head pose parameters; (d) landmark patches.

### 3.2. Sequence-Based Temporal Descriptors Extraction

The facial activity parameters are derived from each frame of video sequences, and they compose of several feature signals along the temporal dimension, including facial configuration (face deformation and head movement) and facial texture. A set of descriptors are extracted from these signals to represent the dynamic features for each video sequence [24].

First, the descriptor signals are temporally smoothed by Butterworth filter (first order with cutoff 1 Hz for temporal signals). Then the first and second derivation of the smoothed signals are estimated for feature description. The original descriptor signal is denoted by  $x$ , the smoothed signal is  $s$ , and its first and second derivation are denoted by  $v$  and  $a$ , respectively. A set of features are extracted from each temporal signal (i.e.,  $s$ ,  $v$  and  $a$ ) to depict the characteristics of the signal variance over time for different aspects of the signal. These parameters include: (1) state parameters: maximum value, minimum value, mean value, and median value; (2) variability parameters: range, standard deviation, inter-quartile range, inter-decile range, and median absolute deviation; (3) peak parameters: instant of time when the amplitude is at its maximum; (4) duration parameters: duration of when the amplitude is greater than mean, and duration of when the amplitude is greater than the average of mean and minimum value; (5) segment parameters: number of segments where the amplitude is greater than mean value, and number of segments where the amplitude is greater than the average of mean and minimum value; (6) area parameters: area between signal and its minimum (AREA), and quotient of AREA (the difference between maximum and minimum value). The first and second derivation of the signals could be treated as the speed and acceleration of the descriptor signals, and these temporal features could describe varying information of facial activity signals, such as amplitude, speed, variability and the time interval of the facial motion. In summary, there are  $16 \times 3$  temporal feature parameters for each facial activity signal.

According to our previous research, the dynamics of temporal facial appearance descriptors are crucial for a description of infant pain expression. LBP-TOP [50] is an effective spatio-temporal histogram representation that captures facial replacement along time. The histogram features are extracted from three orthogonal planes: X-Y (horizontal-vertical spatio), X-T (horizontal spatio-temporal), and Y-T (vertical spatio-temporal). The LBP-TOP is computed on the RAP defined by the 31 facial landmarks, as described in Section 3.1. The dimensions of the dynamic appearance feature for each video is 5487 ( $177 \times 31$ ), which concatenate local appearance descriptors over patches. LBP-TOP based histogram is gray invariant and rotation variant, and the dynamic facial features can tolerate complex environmental conditions to a certain degree.

### 3.3. Dimensionality Reduction Using SLPP

Locality Preserving Projections (LPP) are linear projective maps that approximate the eigenfunctions of the Laplace-Beltrami operator of the manifold. LPP aims at embedding high dimensional features into an intrinsic low-dimensional subspace, which preserves local structure by constructing an adjacency graph denoting the weight matrix. The LPP inherits the advantages of

nonlinear manifold learning and provides transformation function explicitly. The weight matrix is constructed according to the distances between pairs of samples, and the criterion can be applied in an unsupervised or supervised manner. In a supervised manner, the prior class information of the samples is utilized, and the maps will preserve the class structure of the samples. More details of the supervised LPP (SLPP) can be found in [51].

The compact features extracted by SLPP are fed to a classifier to enhance classification performances, and low-dimensional features are mentioned as facial activity features in subsequent sections.

### 3.4. Classification and Decision Fusion

In our paper, we employ Support Vector Machine (SVM) for infant pain recognition. SVM is a well-applied classifier for binary classification and multi-class classification. SVM allows domain-specific selection of the kernel function and can also generalize well with few training data. Each type of facial activity feature trains a SVM classifier individually, and the output is the identification of pain state. Multiple facial activity features are then combined by the decision fusion scheme, i.e., majority voting.

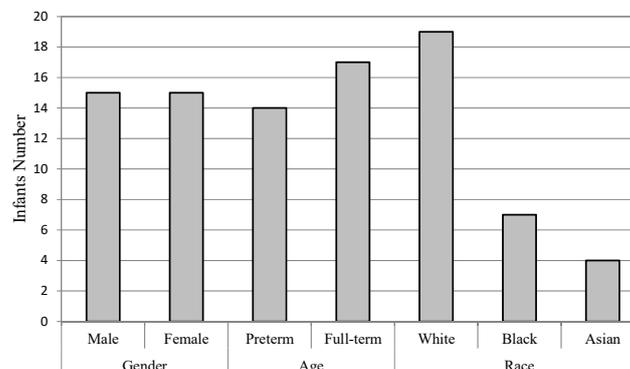
Majority voting is a simple and effective decision fusion method, and fusion criterion is “more than half”. Temporal geometric and texture features are fed to the SVM classifier individually, and each output contributes one vote to the final decision; the major class in the combination is the final label. If there is a tie for different indicators, the class with the highest confidence score is chosen as the final decision of infant pain assessment.

## 4. Dataset

One of the crucial issues in evaluating the performance of a proposed multiple temporal facial activity feature fusion method is the acquisition of infant pain data. To our knowledge, the COPE database [52] is the sole public infant pain database, consisting of 204 static images of 26 infants, face photographed under pain stimuli. However, there is no video data available for research communities. In our experiments, evaluation is conducted on our Infants Pain Assessment Database (IPAD). The detailed information of the IPAD is as follows:

### 4.1. Subjects

A total of 31 infants were recorded during routine painful procedures such as heel lancing, which last for about 5 s, when they were hospitalized in the NICU at Tampa General Hospital. Consent was obtained from the parents of the enrolled infants. Half of the infants were male (50%). The average gestational age of the infants was 36.4 weeks, ranging from 30.4 to 40.6 ( $SD = 2.7$ ). Infants born before 37 gestational weeks are called preterm, and full-term gestation is 37 weeks to 42 weeks. Infants were also racially diverse, with White, Black, and Asian. The distribution of infants' various attributes is shown in Figure 2.



**Figure 2.** Number of infants in different groups.

#### 4.2. Data Acquisition and Preprocessing

The video recordings for infants were acquired by GoPro Hero3+ camera, capturing their facial expression, body movement, and sound. The camera was set in the normal clinical environment, and the videotapes recorded the infants' spontaneous responses during acute painful treatments.

In the preprocessing step, each procedure video is segmented into seven time periods for subsequent analysis. These epochs include five minutes prior to the procedure to provide a baseline state ( $T_0$ ); the pain procedure period ( $T_1$ ), and every minute after the completion of the painful procedure for five minutes ( $T_2 \sim T_6$ ). Each epoch is assigned pain or no-pain, according to the indicator-based pain scale for infants.

#### 4.3. Ground Truth Assessment

The Neonatal Infant Pain Scale (NIPS) [53] is a reliable and valid indicator-based pain scale for both preterm and full-term infants. Both behavioral and physiological indicators are involved, such as facial expression, crying, breathing patterns, arm movements, leg movements, and state of arousal. Each indicator is scored with 0 or 1, except "cry" which has three response categories, i.e., 0, 1, and 2. Total pain score, ranging from 0 to 7, is obtained by summing all the indicator scores. The infants' pain levels are divided into three groups, which are determined by the total pain scores, that is, no pain (0~2), moderate pain (3~4), or severe pain (>4).

The nurses rate the severity of the infants' pain indicators at one-minute intervals during the pain procedure, and the total pain scores are treated as the ground truth for infant pain assessment. The label of pain (3~7)/no-pain (0~2) for each sample is utilized to evaluate the performance of our algorithms.

### 5. Experimental Results and Discussion

The general performance of our automatic infant pain assessment scheme is evaluated in this section. The temporal facial geometric features include facial configuration representation ( $DG_{DisFace}$ ) and head pose representation ( $DG_{DisPose}$ ). The temporal facial texture feature of gradient representation is denoted by  $DA_{Gradient}$ , and LBP-TOP representation is denoted by  $DA_{LBP-TOP}$ . The facial features are low-dimensional data extracted by SLPP and employed to train SVM classifiers individually. The output of the classifier is a binary label of pain or no-pain.

The data set is split into two subsets: one subset is a training set used to develop a classifier model, and the other is a testing set used to evaluate the generalization performance of the classifier model. Due to the limitation of infants' pain samples, it is fair to conduct cross-validation as a powerful general technique. In our experiments, the Leave-One-Subject-Out cross validation is conducted for subject-independent evaluation.

#### 5.1. Comparison of Multi-Feature Fusion

The dynamic geometric features and dynamic appearance features are combined through a decision fusion scheme, and the overall accuracies are compared in Table 1. The differences between recognition accuracies achieved by multiple facial features are identified by a statistical significance test known as *t*-test (Table 2).

Generally, the single dynamic appearance features outperform the single dynamic geometric features, especially when compared to  $DG_{DisPose}$ . The accuracy of dynamic appearance feature  $DA_{LBP-TOP}$  is higher than that of  $DA_{Gradient}$ , and the T-test indicates that there are no significant differences between these two types of facial features. On the other hand, the computational complexity of  $DA_{Gradient}$  is lower than that of  $DA_{LBP-TOP}$ , depending on the original feature dimensions ( $DA_{LBP-TOP}$ :  $5487 = 177 \times 31$ ,  $DA_{Gradient}$ :  $1488 = 16 \times 3 \times 31$ , for each video instance).

Although the performance of a single head pose-related feature is not satisfied, it enhances the accuracy of multi-feature fusion to a certain degree. Difference analysis shows that there are significant differences between single ( $DG_{DisPose}$ ) and multi-features, including two and three feature combinations

( $p < 0.05$ ). It can be seen that the accuracy of a two-feature combination (geometry and appearance) is significantly higher than that of corresponding single features. The three-feature combinations of ( $DG_{DisFace}$  &  $DG_{DisPose}$  &  $DA_{Gradient}$ ) and ( $DG_{DisFace}$  &  $DG_{DisPose}$  &  $DA_{LBP-TOP}$ ) achieve the highest recognition accuracy of 88.48% and 89.33%, respectively. In spite of no significant difference between three-feature combinations and two-feature combinations in the statistical  $t$ -test, the recognition accuracy of ( $DG_{DisFace}$  &  $DG_{DisPose}$  &  $DA_{LBP-TOP}$ ) is approximately more than 2% higher than that of ( $DG_{DisPose}$  &  $DA_{LBP-TOP}$ ). Multiple dynamic facial features provide sufficient information for infant pain assessment and are tolerant to missing data caused by hospitalization environments, such as occlusions. More details of the comparison of multiple dynamic facial features, including feature fusion and decision fusion, are stated in our previous study [48]. Based on these observations, we apply the three-feature decision fusion in subsequent sections for further analysis.

**Table 1.** Comparison of diverse dynamic facial feature combination (%).

	$DG_{DisFace}$	$DG_{DisPose}$	$DA_{Gradient}$	$DA_{LBP-TOP}$	Decision Fusion
Single feature	✓				87.98
		✓			79.97
			✓		85.38
				✓	87.66
Two-feature	✓	✓			87.67
	✓		✓		88.36
		✓	✓		82.55
	✓			✓	87.88
Three-feature		✓		✓	85.51
	✓	✓	✓		88.48
	✓	✓		✓	89.33

**Table 2.** Difference analysis of single facial feature and multiple facial features.

	DG <sub>DisFace</sub>	DG <sub>DisPose</sub>	DA <sub>Gradient</sub>	DA <sub>LBP-TOP</sub>	DG <sub>DisFace</sub> & DG <sub>DisPose</sub>	DG <sub>DisFace</sub> & DA <sub>Gradient</sub>	DG <sub>DisPose</sub> & DA <sub>Gradient</sub>	DG <sub>DisFace</sub> & DA <sub>LBP-TOP</sub>	DG <sub>DisPose</sub> & DA <sub>LBP-TOP</sub>	DG <sub>DisFace</sub> & DG <sub>DisPose</sub> & DA <sub>Gradient</sub>
DG <sub>DisPose</sub>	<b>2.922 (0.004)</b>									
DA <sub>Gradient</sub>	1.345 (0.181)	1.518 (0.131)								
DA <sub>LBP-TOP</sub>	1.028 (0.306)	1.227 (0.222)	0.377 (0.707)							
DG <sub>DisFace</sub> & DG <sub>DisPose</sub>	0.000 (0.999)	<b>2.922 (0.004)</b>	1.345 (0.181)	<b>2.573 (0.011)</b>						
DG <sub>DisFace</sub> & DA <sub>Gradient</sub>	0.446 (0.656)	<b>2.647 (0.009)</b>	1.294 (0.198)	1.958 (0.052)	0.446 (0.656)					
DG <sub>DisPose</sub> & DA <sub>Gradient</sub>	1.000 (0.319)	<b>2.226 (0.027)</b>	0.904 (0.367)	1.214 (0.226)	1.000 (0.319)	1.000 (0.319)				
DG <sub>DisFace</sub> & DA <sub>LBP-TOP</sub>	1.419 (0.158)	<b>2.603 (0.010)</b>	0.894 (0.373)	<b>2.144 (0.033)</b>	1.419 (0.158)	0.446 (0.656)	0.332 (0.740)			
DG <sub>DisPose</sub> & DA <sub>LBP-TOP</sub>	1.345 (0.181)	<b>2.324 (0.021)</b>	0.624 (0.533)	<b>2.264 (0.025)</b>	1.345 (0.181)	0.706 (0.481)	0.000 (0.999)	1.419 (0.158)		
DG <sub>DisFace</sub> & DG <sub>DisPose</sub> & DA <sub>Gradient</sub>	0.000 (0.999)	<b>2.922 (0.004)</b>	1.345 (0.181)	<b>2.573 (0.011)</b>	0.000 (0.999)	0.446 (0.656)	1.000 (0.319)	1.345 (0.181)	0.576 (0.565)	
DG <sub>DisFace</sub> & DG <sub>DisPose</sub> & DA <sub>LBP-TOP</sub>	0.928 (0.355)	<b>2.769 (0.006)</b>	1.728 (0.086)	<b>2.367 (0.019)</b>	0.928 (0.355)	1.096 (0.275)	1.419 (0.158)	0.928 (0.355)	1.351 (0.179)	1.518 (0.131)

Note: The values outside the parentheses are *t*-values, and the values in the parentheses are *p*-values. Significant differences are highlighted in bold (*p* < 0.05).

### 5.2. Pain Assessment in Profile View

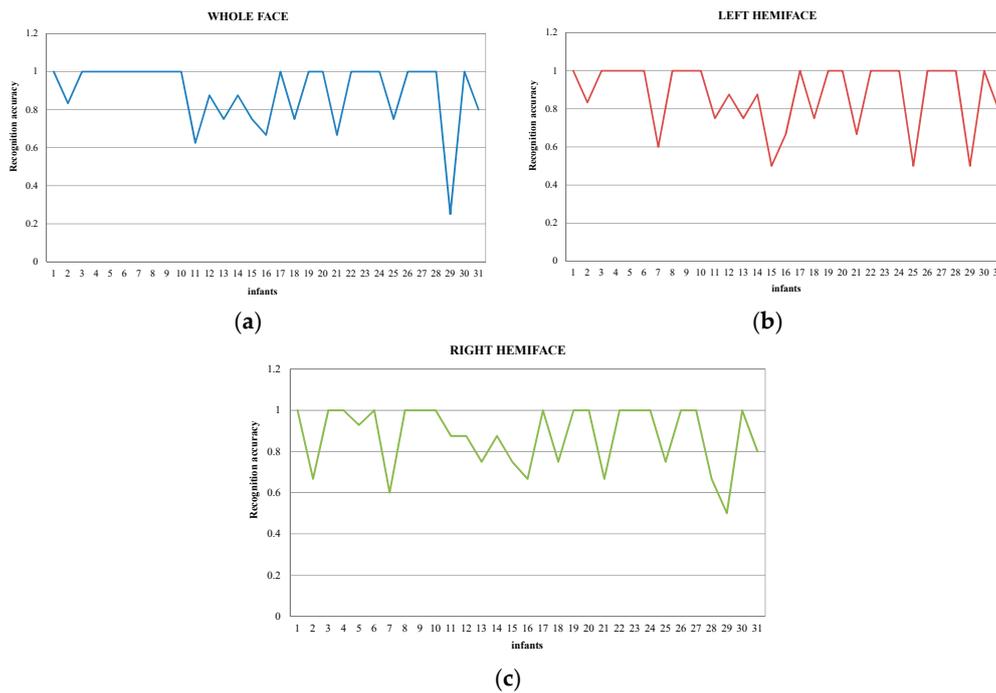
Machine analysis of facial behavior has achieved good recognition accuracies, but they perform only on frontal faces or near-frontal faces, where the individuals face the camera and do not change their head pose three-dimensionally. However, according to clinical observation, head shaking occurs frequently during painful procedures, leading to out-of-plane rotation, which is challenging for pain detection.

In this section, we investigate infant pain assessment by carrying out hemiface facial feature-based fusion and classification. The temporal dynamics of geometric and appearance features of an infant’s left and right side of the face are employed for pain assessment; the results are demonstrated in Figure 3 with confidence intervals denoted by line. The best recognition accuracy is 87.88% of the left hemiface, and is equal to the best recognition rate of the right hemiface, both of which are achieved by (DG<sub>DisFace</sub> & DG<sub>DisPose</sub> & DA<sub>LBP-TOP</sub>). This is consistent with the evidence that spontaneous facial expressions are more symmetrical (involving left hemiface and right hemiface) than deliberate expressions on request [54]. The accuracies of profile evaluation drop by around 1.4%, compared to whole face evaluation. We think it is acceptable, as half of the facial behavior information is missing when hemiface-based pain assessment is applied. The statistical *t*-test indicates that there is no significant difference ( $p < 0.05$ ) between whole face accuracy and hemiface accuracy. It shows that facial symmetry has sufficient discriminating power for infant pain assessment.

Furthermore, the performances of individuals are investigated and the accuracies of (DG<sub>DisFace</sub> & DG<sub>DisPose</sub> & DA<sub>LBP-TOP</sub>) are shown in Figure 4. A majority of the infants achieve accuracy of 100%, while only four infants get low accuracies (<70%) in whole face-based pain recognition. The results of hemiface-based pain recognition are similar, with bad performances for some of the infants. For infants 11, 16, and 21, the no-pain instances are misclassified as pain, due to poor light conditions (dark) and frequent head shaking during a no-pain procedure. The pain instances of infants 29 are misclassified as non-pain, as the infants do not show pain expression during the procedure. As all the infants’ videos are captured in a real NICU, the proposed dynamic facial representation and fusion scheme was able to deal with complex conditions, such as low quality and occlusion.



Figure 3. Overall accuracy comparison of whole face and hemiface model.



**Figure 4.** Individual accuracies of whole face and hemiface models; (a) Whole face (b) Left hemiface (c) Right hemiface.

### 5.3. Individual Variability Analysis

Studies demonstrated that infants’ biobehavioral pain responses vary due to limited neurological development. For a better understanding of the factors that influence pain reactivity in infants, it is necessary to analyze the effects of individual variables, such as gestational age, gender, and race. In this section, the proposed infant pain assessment method is evaluated for these factors. The total IPAD database is divided into diverse groups according to individual variables, and the pain recognition model is constructed for factor-related subsets, respectively (specific model). The experimental results of the specific model are compared to corresponding results of the general model, which is constructed by all the 31 infants. Both overall accuracies and AUC (area under the Receiver Operating Characteristic (ROC) curve) scores are demonstrated in Table 3 for three-feature fusion of ( $DG_{DisFace}$  &  $DG_{DisPose}$  &  $DA_{Gradient}$ ) and ( $DG_{DisFace}$  &  $DG_{DisPose}$  &  $DA_{LBP-TOP}$ ). The best results of overall accuracies are marked in gray, and the best results of AUC scores are marked in green. The experiments are organized into three blocks according to main individual variables: gender, gestational age, and race.

- Gender:** The infants are divided into two groups of male and female. There are 15 infants in each group, as one infant’s gender information is not recorded on the score sheet. For the male group, the overall accuracies of both FDG ( $DG_{DisFace}$  &  $DG_{DisPose}$  &  $DA_{Gradient}$ ) (88.39%) and FDL ( $DG_{DisFace}$  &  $DG_{DisPose}$  &  $DA_{LBP-TOP}$ ) (85.86%) achieved by the specific model are higher than that of the general model (FDG: 82.14%, FDL: 84.52%), while the overall accuracies of female group are just the opposite, i.e., higher accuracies of 94.74% (FDG and FDL) are obtained by the general model. However, no significant differences are found between the specific model and the general model for overall accuracies. Moreover, due to the bias of data distribution, it is also useful to analyze the AUC results. According to the AUC scores, the general model performs better than the specific model. Table 3 shows that although the accuracies of the specific model for the male group are higher, the AUC scores of the general model are much better for both male and female groups, except for the FDL feature in the female group. The ROCs of FDL for the specific model and the general model are illustrated in Figure 5a,b.

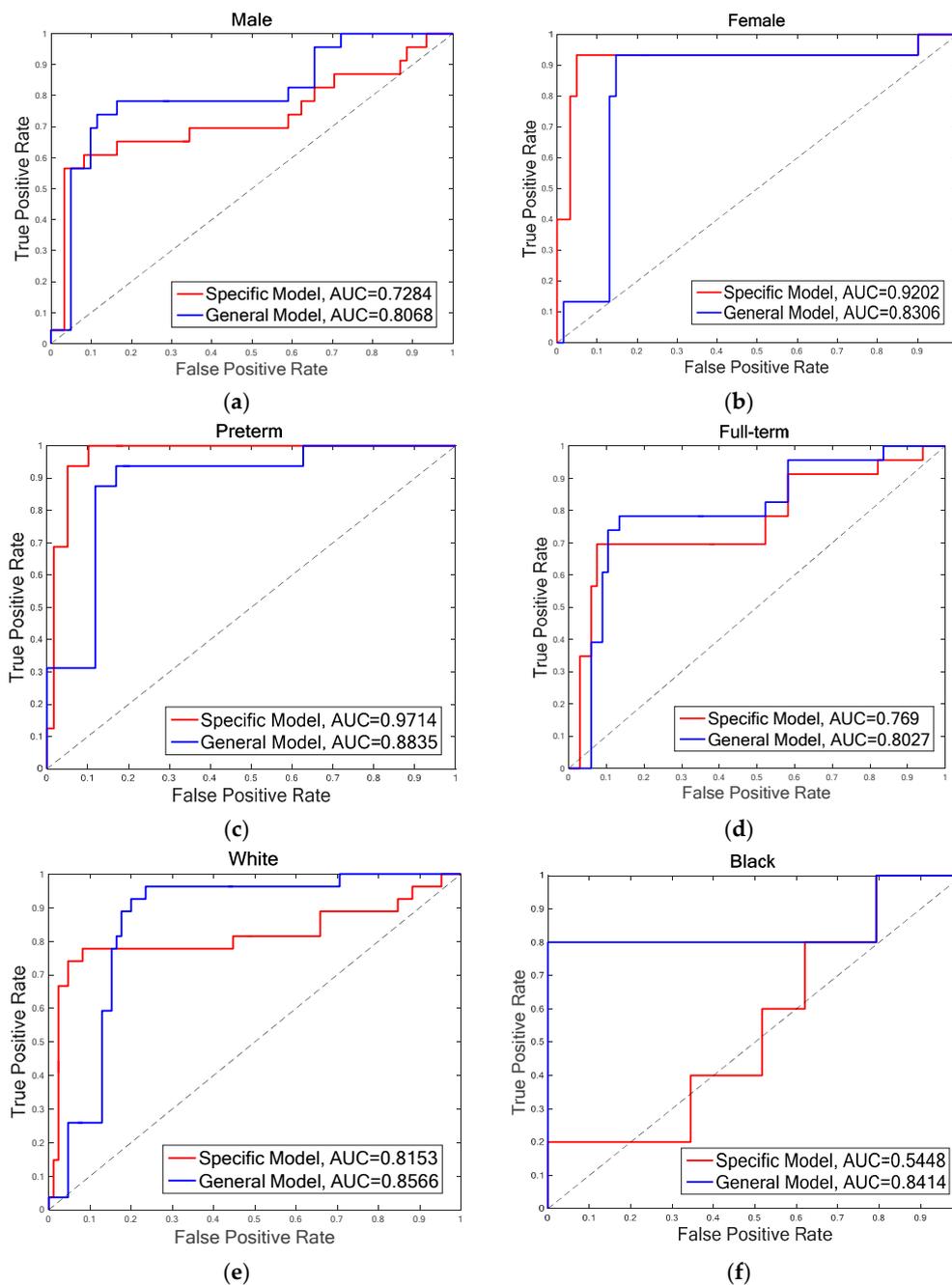
- Gestational age:** The database is divided into two groups of infants with different gestational age, that is, preterm (<37 weeks) and full-term (37 weeks to 42 weeks). For the preterm group, the highest accuracy is 96%, which is achieved by FDG in the specific model and FDL in the general model, while the AUC scores obtained in the specific model are higher than that of the general model for both FDG (0.9894) and FDL (0.9714). The statistical analysis indicates that there is no difference in the other groups in terms of the relationship with the general model. For the full-term group, the highest accuracy (84.44%) and AUC scores (0.8047) are obtained by general model. The general model outperforms the specific model in the aspect of both overall accuracy and AUC score. The ROCs of FDL for two models of the preterm and full-term are shown respectively in Figure 5c,d.
- Race:** The infants belonged to three races: White, Black, and Asian. Only four infants were recorded as Asian; thus, it is not possible to construct an Asian pain assessment model. We examine the effectiveness of specific models for the White group and the Black group, individually. Table 3 demonstrates that the best overall accuracy of the White group is 89.29% (for both FDG and FDL in the general model), and the highest AUC score of 0.9111 is obtained in the general model by FDG. Similar results were found for the Black group; the highest accuracy is 97.06% and AUC score is 0.8414, which was achieved by FDL in the general model. The statistical analysis does not find significant differences between the overall accuracies of the specific model and the general model. More details are demonstrated in Figure 5e,f through the ROCs of diverse models for two infant groups, depending on race.

**Table 3.** Pain assessment results of various factor groups.

			Gender		Age		Race	
			Male	Female	Preterm	Term	White	Black
FDG	Specific	Rate	88.39	89.17	96.00	81.11	89.29	94.12
		AUC	0.7833	0.7738	0.9894	0.7242	0.8436	0.7379
	General	Rate	82.14	94.74	93.33	84.44	89.29	91.18
		AUC	0.8054	0.9224	0.9788	0.8047	0.9111	0.7862
FDL	Specific	Rate	85.86	91.48	93.33	83.33	89.29	88.24
		AUC	0.7284	0.9202	0.9714	0.7690	0.8153	0.5448
	General	Rate	84.52	94.74	96.00	84.44	89.29	97.06
		AUC	0.8068	0.8306	0.8835	0.8027	0.8566	0.8414

Note: FDG means “DG<sub>DisFace</sub> & DG<sub>DisPose</sub> & DA<sub>Gradient</sub>”; FDL means “DG<sub>DisFace</sub> & DG<sub>DisPose</sub> & DA<sub>LBP-TOP</sub>”. AUC means area under Receiver Operating Characteristic (ROC) Curve.

In conclusion, the general model performs better than the specific model for various infant groups of male, female, full-term, white and black. Higher accuracies are obtained in the general model, except for the male group, although no significant difference is found between the two types of models. The best-performed AUC scores were achieved in the general model for the individual variable groups mentioned above. Since the binary labeled infant data is biased, it is valuable to take into consideration the AUC and ROC results. However, for the “preterm” group, better performance is obtained by the specific model for both accuracy and AUC. According to clinical research, infants with low gestational age were not able to express their feelings in the same way as fully developed infants do, and they had fewer behavioral responses. This may be the reason why the specific model is better for preterm infants’ pain detection. Other infant groups do not show such requirements of constructing a pain assessment model by homogeneous data. Therefore, gestational age is the most influencing factor for pain assessment, and it is necessary to construct a specific model depending on gestational age. Besides, it is worth mentioning that the sample size of IPAD (general model) is still small, leading to an even-smaller sample size in specific models. Therefore, it may lead to erroneous conclusions when comparing the specific model to the general model. However, this study provides a pioneer view of the influence factor investigation for automatic infant pain assessment, and we would like to continue collecting data and evaluating our method on a larger dataset.



**Figure 5.** ROC curves for different individual groups with FDL (a) Male (b) Female (c) Preterm (d) Full-term (e) White (f) Black.

## 6. Conclusions

In this study, we propose the dynamics of temporal facial representations of both geometry and appearance; decision fusion is utilized for multi-feature combination. Facial configuration descriptors, head pose descriptors, and gradient descriptors are extracted from the time series of frame-level features; and temporal texture descriptor LBP-TOP is employed to describe facial changes over time. The dynamic facial representation and fusion scheme is successfully applied for infant pain assessment. We also investigate the effects of various factors that influence pain reactivity in infants, such as individual variables of gestational age, gender, and race. Different automatic infant pain assessment models are constructed depending on various influence factors, as well as profile views of faces which

affect the model ability of pain recognition. It can be concluded that the profile-based infant pain assessment is feasible, as its performance is almost as good as the performance of the whole face. The proposed automatic pain assessment scheme is employed on different infant groups, depending on individual variables; and the comparison between the specific model and the general model shows no significant difference between the overall accuracies of these two model types. Moreover, the results of AUC scores are also compared and the general model outperforms the specific model for male, female, full-term, white, and black infant groups. However, for the “preterm” group, better performance is obtained by the specific model for both accuracy and AUC score. Therefore, gestational age is the most influencing factor for infant pain assessment, and it is necessary to construct specific models depending on gestational age. This is mainly because of the limited ability of behavioral communication for infants with low gestational age.

Due to limited sample size of IPAD, some of the conclusions may not be general, when comparing the specific model to the general model, as the former has an even-smaller sample size. On the other hand, there is no control procedure for pain data acquisition. Spontaneous responses are recorded in a normal clinical environment. The labels of the instants are assigned according to a behavioral indicator scale, and it is treated as a gold-standard for infant pain assessment. In terms of future research possibilities, we would like to collect infants’ pain data continuously by taking into consideration a control procedure and evaluate our method on a larger dataset.

**Author Contributions:** Methodology, Experiments, and Writing-Original Draft, R.Z. and G.Z.; Supervision and Writing-Review, D.G. and Y.S.; Data Curation, T.A.

**Funding:** This research was funded by the National Research and Development Major Project (grant number: 2017YFD0400100), the National Natural Science Foundation of China (grant number: 61673052), grant from Chinese Scholarship Council (CSC), and USF Women’s Health Collaborative Grant.

**Acknowledgments:** We are grateful to research coordinators Marcia Kneusel and Judy Zaritt and the entire neonatal staff at Tampa General Hospital for their help and cooperation in data collection. We are especially grateful to the parents who allowed their children to be a part of the study.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Donia, A.E.; Tolba, O.A. Effect of early procedural pain experience on subsequent pain responses among premature infants. *Egypt. Pediatr. Assoc. Gaz.* **2016**, *64*, 74–80. [[CrossRef](#)]
2. Duhn, L.J.; Medves, J.M. A systematic integrative review of infant pain assessment tools. *Adv. Neonatal Care* **2004**, *4*, 126–140. [[CrossRef](#)] [[PubMed](#)]
3. Rojo, R.; Prados-Frutos, J.C.; López-Valverde, A. Pain assessment using the Facial Action Coding System. A systematic review. *Med. Clin.* **2015**, *145*, 350–355. [[CrossRef](#)] [[PubMed](#)]
4. Coghill, R.C.; McHaffie, J.G.; Yen, Y.F. Neural correlates of interindividual differences in the subjective experience of pain. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 8538–8542. [[CrossRef](#)] [[PubMed](#)]
5. Stevens, B.; McGrath, P.; Gibbins, S.; Beyene, J.; Breau, L.; Camfield, C.; Finley, A.; Franck, L.; Howlett, A.; Johnston, C.; et al. Determining behavioural and physiological responses to pain in infants at risk for neurological impairment. *Pain* **2007**, *127*, 94–102. [[CrossRef](#)] [[PubMed](#)]
6. Holsti, L.; Grunau, R.E.; Oberlander, T.F.; Whitfield, M.F. Specific newborn individualized developmental care and assessment program movements are associated with acute pain in preterm infants in the neonatal intensive care unit. *Pediatrics* **2004**, *114*, 65–72. [[CrossRef](#)] [[PubMed](#)]
7. Grunau, R.E.; Haley, D.W.; Whitfield, M.F.; Weinberg, J.; Yu, W.; Thiessen, P. Altered basal cortisol levels at 3, 6, 8 and 18 months in infants born at extremely low gestational age. *J. Pediatr.* **2007**, *150*, 151–156. [[CrossRef](#)] [[PubMed](#)]
8. Slater, R.; Fabrizi, L.; Worley, A.; Meek, J.; Boyd, S.; Fitzgerald, M. Premature infants display increased noxious-evoked neuronal activity in the brain compared to healthy age-matched term-born infants. *Neuroimage* **2010**, *52*, 583–589. [[CrossRef](#)] [[PubMed](#)]
9. Lawrence, J.; Alcock, D.; McGrath, P.; Kay, J.; MacMurray, S.B.; Dulberg, C. The development of a tool to assess neonatal pain. *Neonatal Netw.* **1993**, *12*, 59–66. [[CrossRef](#)]

10. Merkel, S.I.; Voepel-Lewis, T.; Shayevitz, J.R.; Malviya, S. The FLACC: A behavioral scale for scoring postoperative pain in young children. *Pediatr. Nurs.* **1997**, *23*, 293–297. [[PubMed](#)]
11. Grunau, R.E.; Oberlander, T.; Holsti, L.; Whitfield, M.F. Bedside application of the neonatal facial coding system in pain assessment of premature neonates. *Pain* **1998**, *76*, 277–286. [[CrossRef](#)]
12. Hummel, P.; Puchalski, M.; Creech, S.D.; Weiss, M.G. Clinical reliability and validity of the N-PASS: Neonatal pain, agitation and sedation scale with prolonged pain. *J. Perinatal.* **2003**, *28*, 55–60. [[CrossRef](#)] [[PubMed](#)]
13. Debillon, T.; Zupan, V.; Ravault, N.; Magny, J.; Dehan, M.; ABU-SAAD, H. 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*, F36–F41. [[CrossRef](#)] [[PubMed](#)]
14. Krechel, S.W.; Bildner, J. CRIES: A new neonatal postoperative pain measurement score. Initial testing of validity and reliability. *Pediatr. Anesth.* **1995**, *5*, 53–61. [[CrossRef](#)]
15. Sikka, K. Facial expression analysis for estimating pain in clinical settings. In Proceedings of the 16th International Conference on Multimodal Interaction, Istanbul, Turkey, 12–16 November 2014; pp. 349–353.
16. Prkachin, K.M.; Berzins, S.; Mercer, S.R. Encoding and decoding of pain expressions: A judgement study. *Pain* **1994**, *58*, 253–259. [[CrossRef](#)]
17. Riddell, R.P.; Racine, N. Assessing pain in infancy: The caregiver context. *Pain Res. Manag.* **2009**, *14*, 27–32. [[CrossRef](#)]
18. Riddell, R.P.; Horton, R.E.; Hillgrove, J.H.; Craig, K.D. Understanding caregiver judgments of infant pain: Contrasts of parents, nurses and pediatricians. *Pain Res. Manag.* **2008**, *13*, 489–496. [[CrossRef](#)]
19. Bartlett, M.S.; Littlewort, G.C.; Frank, M.G.; Lee, K. Automatic decoding of facial movements reveals deceptive pain expressions. *Curr. Biol.* **2014**, *24*, 738–743. [[CrossRef](#)] [[PubMed](#)]
20. Zhu, S. Pain expression recognition based on pLSA model. *Sci. World J.* **2014**, *2014*, 736106. [[CrossRef](#)] [[PubMed](#)]
21. Ashraf, A.B.; Lucey, S.; Cohn, J.F.; Chen, T.; Ambadar, Z.; Prkachin, K.M.; Solomon, P.E. The painful face-pain expression recognition using active appearance models. *Image Vis. Comput.* **2009**, *27*, 1788–1796. [[CrossRef](#)] [[PubMed](#)]
22. Littlewort, G.C.; Bartlett, M.S.; Lee, K. Faces of pain: Automated measurement of spontaneous facial expressions of genuine and posed pain. In Proceedings of the 9th International Conference on Multimodal Interfaces, Nagoya, Japan, 12–15 November 2007; pp. 15–21.
23. Hammal, Z.; Cohn, J.F. Automatic detection of pain intensity. In Proceedings of the ACM International Conference on Multimodal Interaction (ICMI), Santa Monica, CA, USA, 22–26 October 2012; Volume 7325, pp. 47–52.
24. Werner, P.; Al-Hamadi, A.; Niese, R.; Walter, S.; Gruss, S.; Traue, H.C. Towards pain monitoring: Facial expression, head pose, a new database, an automatic system and remaining challenges. In Proceedings of the British Machine Vision Conference, Bristol, UK, 9–13 September 2013.
25. Paraschiv-Ionescu, A.; Buchser, E.E.; Rutschmann, B.; Najafi, B.; Aminian, K. Ambulatory system for the quantitative and qualitative analysis of gait and posture in chronic pain patients treated with spinal cord stimulation. *Gait Posture* **2004**, *20*, 113–125. [[CrossRef](#)] [[PubMed](#)]
26. Scheiner, E.; Hammerschmidt, K.; Jürgens, U.; Zwirner, P. Acoustic analyses of developmental changes and emotional expression in the preverbal vocalizations of infants. *J. Voice* **2002**, *16*, 509–529. [[CrossRef](#)]
27. El Ayadi, M.; Kamel, M.S.; Karray, F. Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognit.* **2011**, *44*, 572–587. [[CrossRef](#)]
28. Yang, Y.; Fairbairn, C.; Cohn, J.F. Detecting depression severity from vocal prosody. *IEEE Trans. Affect. Comput.* **2013**, *4*, 142–150. [[CrossRef](#)] [[PubMed](#)]
29. Johnston, C.C.; Strada, M.E. Acute pain response in infants: A multidimensional description. *Pain* **1986**, *24*, 373–382. [[CrossRef](#)]
30. Craig, K.D.; Grunau, R.V.; Aquan-Assee, J. Judgement of pain in new-borns: Facial activity and cry as determinants. *Can. J. Behav. Sci.* **1988**, *20*, 442–451. [[CrossRef](#)]
31. Sariyanidi, E.; Gunes, H.; Cavallaro, A. Automatic analysis of facial affect: A survey of registration, representation, and recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1113–1133. [[CrossRef](#)] [[PubMed](#)]
32. Kaltwang, S.; Rudovic, O.; Pantic, M. Continuous pain intensity estimation from facial expressions. *Adv. Vis. Comput.* **2012**, *7432*, 368–377.
33. Lo Presti, L.; La Cascia, M. Boosting hankel matrices for face emotion recognition and pain detection. *Comput. Vis. Image Underst.* **2016**, *156*, 19–33. [[CrossRef](#)]

34. Rudovic, O.; Pavlovic, V.; Pantic, M. Automatic pain intensity estimation with heteroscedastic conditional ordinal random fields. In Proceedings of the International Symposium on Visual Computing, Rethymnon, Crete, Greece, 29–31 July 2013; Volume 8034, pp. 234–243.
35. Oster, H. *Baby FACS: Facial Action Coding System for Infants and Young Children*; New York University: New York, NY, USA, 2007.
36. Peters, J.W.; Koot, H.M.; Grunau, R.E.; de Boer, J.; van Druenen, M.J.; Tibboel, D.; Duivenvoorden, H.J. Neonatal facial coding system for assessing postoperative pain in infants: Item reduction is valid and feasible. *Clin. J. Pain* **2003**, *19*, 353–363. [[CrossRef](#)] [[PubMed](#)]
37. Brahnam, S.; Chuang, C.F.; Shih, F.Y.; Slack, M.R. SVM classification of neonatal facial images of pain. In *Fuzzy Logic and Applications*; Bloch, I., Petrosino, A., Tettamanzi, A.G.B., Eds.; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2005; Volume 3849, pp. 121–128.
38. Brahnam, S.; Chuang, C.F.; Sexton, R.S.; Shih, F.Y. Machine assessment of neonatal facial expressions of acute pain. *Decis. Support Syst.* **2007**, *43*, 1242–1254. [[CrossRef](#)]
39. Gholami, B.; Haddad, W.M.; Tannenbaum, A.R. Agitation and pain assessment using digital imaging. In Proceedings of the Annual International Conference of the IEEE in Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 2–6 September 2009; pp. 2176–2179.
40. Nanni, L.; Brahnam, S.; Lumini, A. A local approach based on a Local Binary Patterns variant texture descriptor for classifying pain states. *Expert Syst. Appl.* **2010**, *37*, 7888–7894. [[CrossRef](#)]
41. Fotiadou, E.; Zinger, S.; Tjon a Ten, W.E.; Oetomo, S.B.; de With, P.H.N. Video-based facial discomfort analysis for infants. *Vis. Inf. Process. Commun. V* **2014**, *9029*, 90290F.
42. Zamzmi, G.; Pai, C.; Goldgof, D.; Kasturi, R.; Ashmeade, T.; Sun, Y. An approach for automated multimodal analysis of infants' pain. In Proceedings of the 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 4143–4148.
43. Zamzmi, G.; Pai, C.; Godgof, D.; Kasturi, R.; Sun, Y.; Ashmeade, T. Automated pain assessment in neonates. Scandinavian Conference on Image Analysis, Tromsø, Norway, 12–14 June 2017; pp. 350–361.
44. Pantic, M.; Rothkrantz, L.J.M. Facial action recognition for facial expression analysis from static face images. *IEEE Trans. Syst. Man Cybern. Part B* **2004**, *34*, 1449–1461. [[CrossRef](#)]
45. Pantic, M.; Patras, I. Dynamics of facial expression: Recognition of facial actions and their temporal segments from face profile image sequences. *IEEE Trans. Syst. Man Cybern. Part B* **2006**, *36*, 433–449. [[CrossRef](#)]
46. Valeri, B.O.; Linhares, M.B.M. Pain in preterm infants: Effects of sex, gestational age, and neonatal illness severity. *Psychol. Neurosci.* **2012**, *5*, 11–19. [[CrossRef](#)]
47. Derbyshire, S.W.G. Gender, pain and the brain. *Pain Clin. Updates* **2008**, *16*, 1–4.
48. Zhi, R.C.; Zamzmi, G.; Goldgof, D.; Ashmeade, T.; Sun, Y. Infants' pain recognition based on facial expression: Dynamic hybrid descriptions. *IEICE Trans. Inf. Commun.* **2018**. [[CrossRef](#)]
49. Cootes, T.F.; Edwards, G.J.; Taylor, C.J. Active appearance models. In *Lecture Notes in Computer Science, Proceedings of the 5th European Conference on Computer Vision Freiburg, Freiburg, Germany, 2–6 June 1998*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 484–498.
50. Zhao, G.; Pietikäinen, M. Texture recognition using local binary patterns with an application to facial expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 915–928. [[CrossRef](#)] [[PubMed](#)]
51. Shermira, J. Application of Locality Preserving Projections in face recognition. *Int. J. Adv. Comput. Sci. Appl.* **2010**, *1*, 82–85.
52. Brahnam, S.; Chuang, C.F.; Shih, F.Y.; Slack, M.R. Machine recognition and representation of neonatal facial displays of acute pain. *Artif. Intell. Med.* **2006**, *36*, 211–222. [[CrossRef](#)] [[PubMed](#)]
53. Hudson-Barr, D.; Capper-Michel, B.; Lambert, S.; Palermo, T.M.; 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. [[CrossRef](#)] [[PubMed](#)]
54. Hager, J.C. Asymmetry in facial muscular actions. In *What the Face Reveals*; Ekman, P., Rosenberg, E.L., Eds.; Oxford University Press: New York, NY, USA, 1997; pp. 58–62.

