Emotion detection in infants using morphological operations

|  |  |  |
| --- | --- | --- |
| Prathibha H M | Dr N Sriraam | Tejaswini S |
| MSRIT | MSRIT | MSRIT |
| [prathibha@msrit.edu](mailto:prathibha@msrit.edu) | sriraam@msrit.edu | tejaswini.s@msrit.edu |

***Abstract*— Infants with different health conditions are monitored at NICU. Infants rely on nonverbal communication, hence monitoring the emotions are important. In this context this study aims at detecting different emotions like sleeping, crying and neutral faces using morphological operations. Infant videos are collected from NICU, frames are extracted from these videos and preprocessing is performed. Face region is selected by cropping manually and non-skin region is extracted. Morphological operations are performed to extract the boundaries of eye and mouth region based on which emotion is detected. the results of statistical analysis of ANOVA performed for hunger, discomfort and pain and students t-test was performed for different features like cry vs sleep, cry vs neutral and sleep vs neutral. It was observed that all features were significant except sleep vs neutral.**

Keywords—Facial features, Emotion recognition, segmentation, morphological operation

# Introduction

# New-born babies who need intensive medical care are often put in a special area of the hospital called the neonatal intensive care unit (NICU). Babies which are born preterm or with any health disorders are placed in NICU to provide special care and monitor continuously. New-born infants are exposed to painful experiences that might increase their short- and long-term morbidity and mortality, in addition to being associated with neurological developmental disorders [1].Pain is an unpleasant emotion associated with actual or potential tissue damage. Accurate assessment of pain is vital to ensure the optimal effectiveness and safety of pain management therapy in neonates who experience pain during the course of their NICU stay [2]. Accurate pain measurements in infants are difficult to achieve. Pain in infants are measured through behavioral changes such as crying, facial expressions, body posture, and movements. These behavior depends on infant’s gestational age, and maturity. Cry characteristics are also not good indicators in preterm or acutely ill infants, as it is difficult for them to produce a robust cry. Pain scales such as neonatal infant pain scale (NIPS); neonatal facial coding system (NFCS); neonatal pain, agitation, and sedation scale(N-PASS); cry, required oxygen, increased vital signs, expression, sleeplessness scale (CRIES); COMFORT Scale; and Douleur Aigue Nouveau-ne (DAN) scoring system are used for pain validation.[3]. Infants cannot express pain verbally, so this impossibility has created the necessity of using other media for its evaluation and detection. In this way, pain scales based on vital signals and facial changes have been created to evaluate the pain of neonates. [4]

Facial expression recognition system identifies different facial features or landmarks are statistically related with pain stimulus. E. Fotiadou et.al. proposed a study based on geometrical features of face, Active appearance model (AAM) which tracks the newborn’s face and its features. Here texture based descriptors are employed to detect the discomfort using SVM classifier. [5] .Yue Sun et al. proposed a study based on discomfort detection. The study is being divided into 2 phases 1st phase derives geometrical and appearance parameters and 2nd phase includes facial landmark detection. SVM classifier is used recognize facial expression of discomfort.[6]. Wenping Lu et al. proposed a study based on the analysis of infants geometrical (eye and mouth) features for facial expression recognition and categorizing into crying, happy, sleepy and normal classes.[7]. Chiung-Yao Fang et al. proposed a study for emotion recognition system which detects facial expression based on infant skin color by deploying a locus model. Principal component analysis is use to extract features and classify them into sleeping, dazing, crying, laughing, yawning, sneezing and vomiting.[8] . Ruicong Zhi et al proposed a study for the assessment of infant pain based on dynamic pain facial expressions and fusion scheme for automatic pain assessment in infants by combining temporal appearance facial features and temporal geometric facial features. The effects of various factors that influence pain reactivity in infants, such as individual variables of gestational age, gender, and race are investigated. SVM classifier is used for infant pain recognition. [9]

# Methodology

The proposed work is evaluation of the existing algorithm proposed by A. Mercy Rani et, al [9] for infant emotion detection. Infant’s face is detected from the frames extracted from collected video. Selection of frames is based on the Structural similarity index(SSIM), Peak signal to noise ratio values(PSNR) and selection of Random frames. Preprocessing is performed to remove noise present in the image and filtering is applied. Selection of ROI i.e., face region is done by manual cropping. Non skin regions are extracted by thresholding using RGB color space. Morphological operation is done in order to obtain the boundaries of eye and mouth regions. Emotion recognition is performed by calculating the area of eye and mouth region. Student t-test and ANOVA is performed.

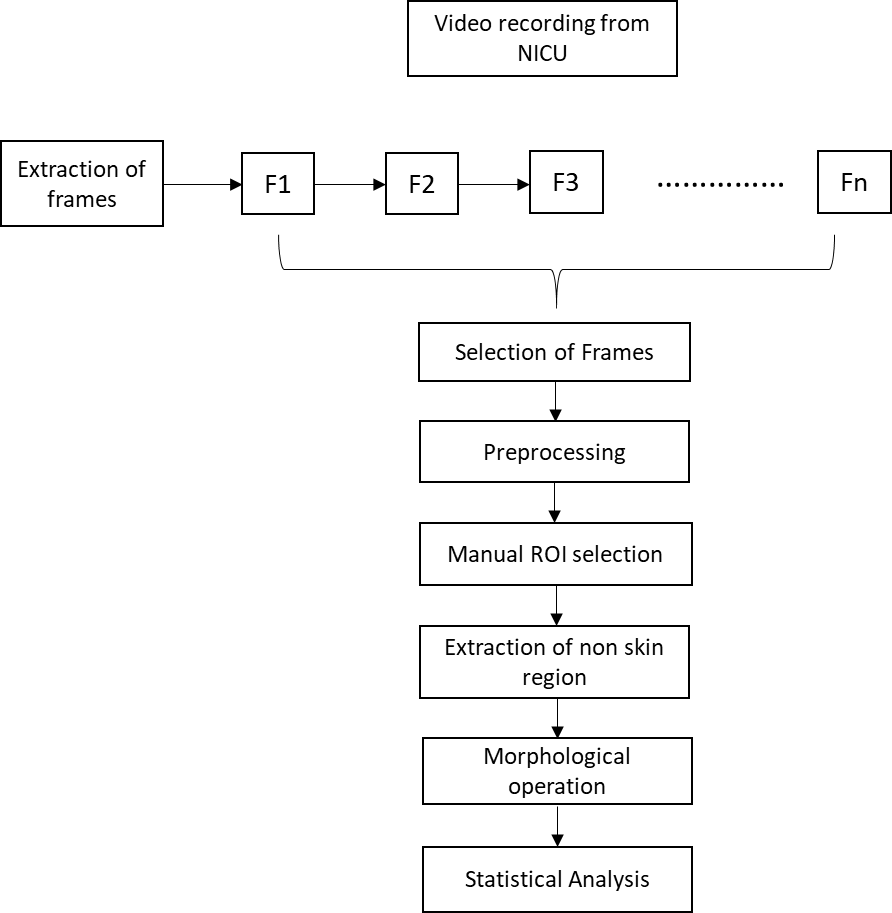


Figure 1 Proposed work flow

## Database

The images selected for analysis are collected from NICU at MS Ramaiah Teaching Hospital, Bengaluru. Different activities of the baby are recorded which includes cries and sleep of the baby using Fingers 1080 Hi-Res Webcam. NICU includes both full term and preterm neonates with gestation age ranging from 28-37 weeks and with varied health conditions. Different activities of the baby are recorded which includes cries and sleep of the baby using Fingers 1080 Hi-Res Webcam. Out of 200 videos of neonates ,130 videos are discarded due to exclusion criteria listed below.70 videos are available for analysis. Videos are recorded at a frame rate of 30 fps. Camera was placed beside the infant warmer focusing the baby face area with the help of a customized stand. The stand is undisturbed during the process of recording as shown in Figure 2.

For facial expression analysis, videos are first converted into frames. 20% of frames are taken for analysis. These frames are then selected based on following criteria

(a) Selecting Random frames -20% of frames are taken for analysis by considering first, mid and last frame.

(b) Selecting frames based on SSIM (Structural Similarity Index) values by varying threshold values.

(c) Selection based on correlation and PSNR values.

* Exclusion criteria

Not all videos recorded at NICU are considered for further processing some videos are excluded for following reasons

* + Baby turning to side where baby face is not focused by camera.
  + Baby face covered by ventilators or baby under CPAP device (Continuous positive airway pressure).
  + Baby under light therapy.
  + Improper lighting conditions.
  + Disturbance in camera set-up during recording which changes the focus.

Below table provides the details of recordings which includes gender, gestation age, duration of recording, frame rate, diagnosis details and no: of still frames selected.



Figure 2 Camera set-up at NICU

Table 1 Database details

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Patient Id | Gender | Gestation week(weeks) | Duration  (Min) | Frame rate | Camera used | Diagnosis | No: of still frames selected |
| PTV001 | M | 34 | 04.09 | 30fps | Wifi cam | PT | 10 |
| PTV017 | M | 37 | 01.10 | 30fps | Mobile | PT/AGA | 10 |
| PTV018 | M | 28 | 03.39 | 30fps | Mobile | PT/AGA/RDS | 10 |
| PTV016 | M | 32 | 21.55 | 30fps | Wifi cam | PT | 10 |
| PTV015 | M | 32 | 03.00 | 30fps | Mobile | PT | 10 |
| FTV003 | M | 40 | 03.00 | 30fps | Mobile | SGA | 10 |
| FTV001 | M | 37 | 21.96 | 30fps | Wifi cam | AGA/MAS | 10 |
| FTV002 | M | 40 | 21.55 | 30fps | Wifi cam | AGA/ MSAF | 10 |
| PTV006 | F | 33 | 21.55 | 30fps | Wifi cam | PT/AGA | 10 |
| PTV007 | M | 33 | .21.55 | 30fps | Wifi cam | PT/AGA/Rh-ve mother | 10 |
| PTV008 | M | 33 | 21.55 | 30fps | Wifi cam | PT/AGA | 10 |
| PTV009 | F | 34 | 21.55 | 30fps | Wifi cam | PT / AGA | 10 |
| PTV010 | F | 34 | 21.55 | 30fps | Wifi cam | PT / AGA | 10 |
| PTV011 | M | 36 | 21.55 | 30fps | Wifi cam | Late PT | 10 |
| PTV012 | M | 36 | 21.55 | 30fps | Wifi cam | PT/Perinatal depression | 10 |
| PTV013 | F | 29 | 21.55 | 30fps | Wifi cam | Early PT/AGA | 10 |
| PTV002 | M | 34 | 21.96 | 30fps | Wifi cam | PT/AGA/IDM | 10 |

**Diagnosis details**

* PT – Pre term
* AGA -Appropriate for Gestational Age.
* RDS- Respiratory distress syndrome
* SGA- Small for gestational age
* MAS - Meconium aspiration syndrome
* MSAF -Meconium-stained amniotic fluid
* IDM- Infant of diabetic mother.



Figure 3 Input Frames

Figure 2 shows the image selected for emotion detection. Emotion classification is mainly done based on 3 conditions,

1. Neutral - Eye open and mouth shut.
2. Crying – Eye shut and mouth open.
3. Sleep – Both eyes and mouth are shut.

## Preprocessing

This stage includes the removal of noise present in the image and improve the quality of image. Median filter is used for noise removal. The quality of the filtered image is measured using PSNR (Peak signal-to-noise ratio) and RMSE (Root mean square error) values.

## Manual Face Cropping

The region of interest is the baby face. Since most of the face detection algorithm fails to detect baby’s face here we crop baby’s manually. Grid size is different for each babies because of variation in baby head positioning and camera placement.



Figure 4 Cropped face

## Extraction of non skin region

Skin tone is one of distinguishable feature of human face. Emotions can be identified from non-skin regions of face. Extraction of non-skin regions i.e., eyes and mouth from face image using RGB color space. Thresholding technique is applied to RGB image to extract non-skin pixels.

R > 95 && G > 40 && B > 20



Figure 5 Segmented image

## Morphological operation

Morphological processing is carried out to refine skin regions extracted from segmentation and to extract geometrical features of facial region reducing the noise and unwanted region. Morphological opening of the image, erosion followed by boundary extraction to obtain eyes and mouth regions accurately.

Erosion removes small-scale details from a binary image but simultaneously reduces the size of regions of interest. By subtracting the eroded image from the original image, boundaries of each region can be found.

The assigned structuring element is used for probing and expanding the shapes contained in the input image. It acts like local maximum filter. Dilation has the opposite effect to erosion. It adds a layer of pixels to both the inner and outer boundaries of regions.

Opening is so called because it can open up a gap between objects connected by a thin bridge of pixels. Any regions that have survived the erosion are restored to their original size by the dilation.

## Emotion recognition

Emotion recognition is done by calculating the area around the mouth region. Area around mouth region is extracted using area open function and filling any holes inside the region of interest. The area of the mouth region is calculated based on the number of pixels multiplied with the pixel width.



Figure 6 Extracted Mouth region

## Statistical Ananlysis

The significance value of each method were compared with neutral, sleeping and crying states. It was noted that all the methods were found to be significant (p<0.05). Table 2 presents ANOVA test results where P1-P5 are threshold values. Table 3 presents Student’s t test results test results where P1-P5 are threshold values.

Table 2 ANOVA test results

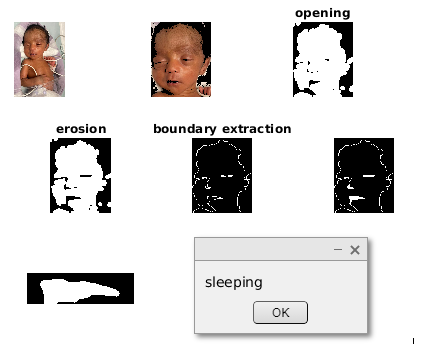
|  |  |
| --- | --- |
| **Features** | **ANOVA** |
| P1 | 0.00 |
| P2 | 0.01 |
| P3 | 0.00 |
| P4 | 0.07 |
| P5 | 0.03 |

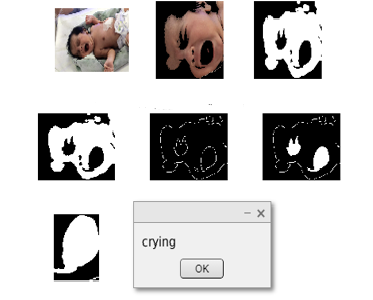
Table 3 Student's t test results

|  |  |  |  |
| --- | --- | --- | --- |
| **Features** | **Cry vs Sleep** | **Cry vs Neutral** | **Sleep vs Neutral** |
| P1 | 0.00 | 0.00 | 0.01 |
| P2 | 0.00 | 0.02 | 0.12 |
| P3 | 0.00 | 0.00 | 0.31 |
| P4 | 0.01 | 0.00 | 0.31 |
| P5 | 0.00 | 0.10 | 0.17 |

# RESULTS

Videos of babies are recorded from NICU. Frames are extracted from these videos; these frames are considered as input image. Filtering is applied to remove noise from the image and ROI is detected. Segmentation is applied to extract the non-skin regions from the baby face. Morphological operations are carried out to extract the geometrical boundaries of mouth and eyes region. Emotion recognition is based on the mouth region.





# DISCUSSION AND CONCULSION

Videos from NICU are recorded focusing baby face. Frames are extracted from these videos using SSIM, PSNR and random frame selection methods. Noise removal from these images is performed at preprocessing stage. Face detection and ROI selection is done by cropping the image manually. Image segmentation is performed to extract non-skin regions like eyes and mouth using thresholding technique. Morphological processing is performed to refine extracted regions. Boundaries around mouth and eye region are extracted and area is calculated and threshold is set to detect emotion from infant face. Statistical analysis is performed and Table 2 and 3 shows the results of statistical analysis of ANOVA performed for hunger, discomfort and pain and students t-test was performed for different features like cry vs sleep, cry vs neutral and sleep vs neutral. It was observed that all features were significant except sleep vs neutral.

Our future work aims at working with increased sample size, extraction of other facial regions apart from eyes and nose to detected other emotions from infant face and automatic face detection to avoid manual intervention.

##### References

[1] T. M. Heiderich, A. Teresa, F. Stochero, and R. Guinsburg, “Neonatal procedural pain can be assessed by computer software that has good sensitivity and specificity to detect facial movements,” pp. 63–69, 2015, doi: 10.1111/apa.12861.

[2] L. G. Maxwell, M. V Fraga, and C. P. Malavolta, “Assessment of Pain in the N e w b o r n : An Update Neonate Pain Assessment Pain scales,” vol. 46, pp. 693–707, 2019, doi: 10.1016/j.clp.2019.08.005.

[3] R. Srouji, S. Ratnapalan, and S. Schneeweiss, “Pain in Children : Assessment and Nonpharmacological Management,” vol. 2010, 2010, doi: 10.1155/2010/474838.

[4] A. Martínez, F. A. Pujol, and H. Mora, “Application of texture descriptors to facial emotion recognition in infants,” *Appl. Sci.*, vol. 10, no. 3, 2020, doi: 10.3390/app10031115.

[5] E. Fotiadou, S. Zinger, W. E. Tjon, S. B. Oetomo, and P. H. N. De With, “Video-based facial discomfort analysis for infants,” vol. 9029, pp. 1–14, 2014, doi: 10.1117/12.2037661.

[6] Y. Sun *et al.*, “Video-based discomfort detection for infants,” *Mach. Vis. Appl.*, vol. 30, no. 5, pp. 933–944, 2019, doi: 10.1007/s00138-018-0968-1.

[7] L. Wenping, L. Min, H. Wenfei, and L. Yu, “A Facial Expression Recognition Method for Baby Video Surveillance,” *Proc. 3rd Int. Conf. Multimed. Technol.*, vol. 84, pp. 516–523, 2013, doi: 10.2991/icmt-13.2013.63.

[8] I. Engineering, “AN INFANT FACIAL EXPRESSION RECOGNITION SYSTEM.”

[9] jeetu, “F5131086617,” no. 6, pp. 101–106, 2017.

[10] R. Zhi, G. Zamzmi, D. Goldgof, T. Ashmeade, and Y. Sun, “Automatic Infants’ Pain Assessment by Dynamic Facial Representation: Effects of Profile View, Gestational Age, Gender , and Race,” pp. 1–16, doi: 10.3390/jcm7070173.