**Detecting facial features which includes face detection, non-skin region extraction and morphological processing**

**(**Image Processing Techniques To Recognize Facial Emotions, A. Mercy Rani et,al 2017)

The images selected for analysis are taken from online sources and images collected from MSRMH.

A picture containing person

Description automatically generatedA person wearing goggles

Description automatically generated with low confidenceA picture containing indoor, laying, bed, baby

Description automatically generatedA baby in a crib

Description automatically generated with medium confidence

A picture containing baby, person, laying, bed

Description automatically generatedA baby crying with its mouth open

Description automatically generated with medium confidenceA baby with his mouth open

Description automatically generated with low confidenceA baby lying on its back

Description automatically generated with low confidence

A picture containing person, indoor, wall, baby

Description automatically generatedA picture containing baby, person, indoor, high

Description automatically generated

Analysis is done using MATLAB. Algorithm is as mentioned below,

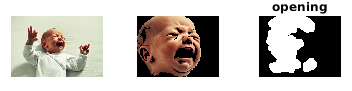
* Read the input video frame image
* Convert the image into grayscale image.
* Enhance the input image with median, wiener and Gaussian filters.
* Find the best filter based on PSNR, RMSE values.
* Apply viola-jones algorithm to detect the face region.
* Use bounding box method and crop the face region.
* Use threshold value to extract non skin regions.
* Apply morphological operations to extract continuous boundaries of non-skin region.
* Mask the boundary from the original image.
* Extract the mouth region.
* Area is calculated from extracted mouth region.
* Recognize facial emotions based on the value of area.

**Outputs**

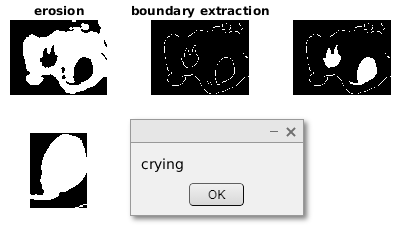
   

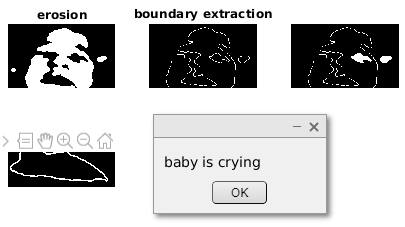
 

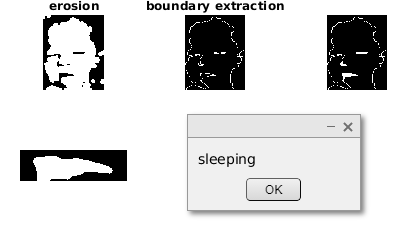






**Results**

Infant face images are extracted from recorded video from NICU and downloaded from online sources. Viola johns algorithm is used to detect baby face. The images taken in this study consists of preterm and full-term babies and facial features of babies are not well distinguished. Hence Viola john’s algorithm fails to detect baby face. Instead of that crop feature can be used. Algorithm classifies images as happy, neutral, sleeping and crying categories. Algorithm needs to be tested on a greater number of images for better accuracy. With the available data it is found that the proposed algorithm works well for images recorded from mobile camera.

**Infant facial expression recognition**

(Facial Expression Recognition Method for Baby Video Surveillance, Wenping Lu et,al 2013)

The images selected for analysis are recorded from NICU at MSRTH using Wi-Fi camera and mobile camera. 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. Videos are recorded at a frame rate of 30fps.

**Table 1 Dataset details**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **PATIENT ID** | **Gender** | **Gestation Week** | **Duration of recording** | **Frame Rate** | **Camera used for recording** | **Diagnosis** |
| PTV001 | M | 34 weeks | 00.04.09 | 30fps | Wifi camera | PT |
| PTV017 | M | 37 weeks | 00.01.10 | 30fps | Mobile | PT/AGA/M |
| PTV018 | M | 28 weeks | 00.03.39 | 30fps | Mobile | PT/AGA/RDS |
| PTV016 | M | 32 weeks | 00.21.55 | 30fps | Wifi camera | Preterm / Perinatal Depression |
| PTV015 | M | 32 weeks | 00.03.00 | 30fps | Mobile | Late PT/AGA |
| FTV003 | M | 40 weeks | 00.03.00 | 30fps | Mobile | Term/SGA/Male/Congenital Hydrocephalus |
| FTV001 | M | 37 Weeks | 00.21.96 | 30fps | Wifi camera | Term /AGA/ MAS |
| FTV002 | M | 40 weeks | 00.21.55 | 30fps | Wifi camera | Term /AGA/ MSAF |

 A picture containing person, cellphone, ear

Description automatically generated A baby in a crib

Description automatically generated with medium confidence A person wearing goggles

Description automatically generated with low confidence

A picture containing indoor, laying, bed, baby

Description automatically generated A baby wearing sunglasses

Description automatically generated with low confidence A picture containing person

Description automatically generated A picture containing person, indoor

Description automatically generated

Images of babies used.

Video input

Conversion of video to frames

Pre-processing

Binarization

Emotion recognition

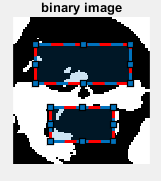
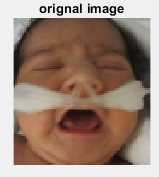
Calculate eye and mouth region.

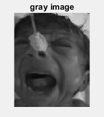
**Proposed block diagram**

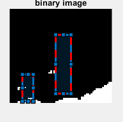
Analysis is done using MATLAB. Algorithm is as mentioned below,

1. Input video and convert into frames.
2. Detect the infant face in the image and crop the region off the image,.
3. Resize cropped image. RGB to gray conversion.
4. Convert the gray into binary image .
5. Calculate area around eye and mouth (D)
6. If ( D<=-200 )
7. the facial expression is “Crying”
8. Elseif ( Se<=100 && Sm<=100 && D<=30 && D>=-70 )
9. the facial expression is “Sleeping”
10. Elseif ( D>=0 )
11. the facial expression is “Normal”
12. end

**Output**

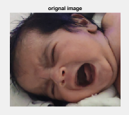
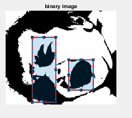


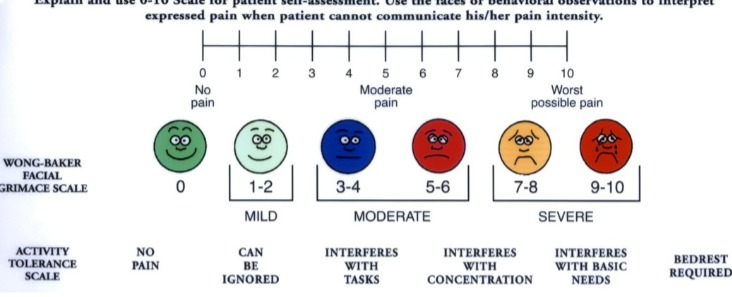
  

Results

The database of infant face is downloaded from online sources which are taken at a fixed camera position. the images which are considered in this evaluation are derived from videos recorded from NICU. Due to the movement of baby face localization is a challenging task to calculate the above-mentioned parameters.

**Classification results**

* Support vector machine (SVM) is used for classification.
* Video data collected from NICU at MSRH are annotated and labeled into three classes i.e., Moderate, None and Severe pain based on the Wong bakers pain scale.



* Labelled data is divided into training and testing data.
* Moderate, none and severe are the three classes.



* Some of the key parameters in SVM are:

→Gamma : defines how far the influence of single training examples reaches values leads to biased results.

→ C : Controls the cost of miscalculations

Small C — makes the cost of misclassification LOW

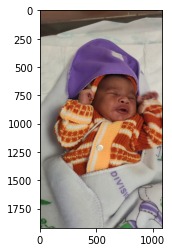
Large C — makes the cost of misclassification HIGH

→ Kernel : SVM algorithms use a set of mathematical functions that are defined as the kernel.

* RBF(Radial Basis Function), Polynomial Kernel are deployed.



Below are results of classifier

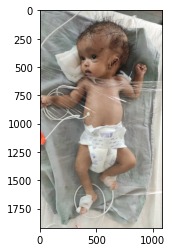


**moderate = 95.3253408163106%**

**None = 2.420899909551253%**

**severe = 2.2537592741381443%**

**The predicted image is : moderate**

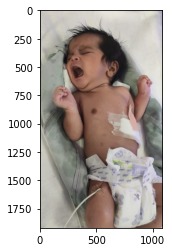


**moderate = 25.095100341583905%**

**None = 60.30653489330289%**

**severe = 14.598364765113189%**

**The predicted image is : None**

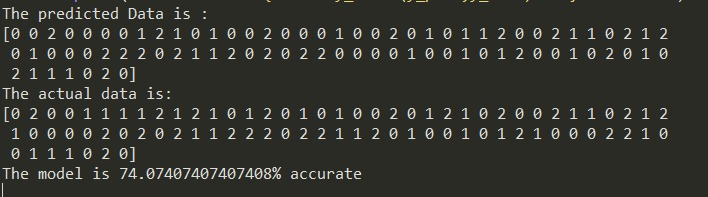


**moderate = 8.648209596205698%**

**None = 3.2216067411114846%**

**severe = 88.13018366268282%**

**The predicted image is: Severe**



* Model is also tested for completely unknown data



**moderate = 40.38214589165237%**

**None = 20.109537893228474%**

**severe = 39.50831621511915%**

**The predicted image is : moderate**

**recall score for all three classes (0.7773809523809524, 0.7374162381175005, 0.7483076755104269)**

* Algorithm gives **74%** accuracy.

**Drawbacks**

* SVM algorithm is not suitable for large data sets.
* If the data is irregular, accuracy drops significantly.
* If the number of features for each data point exceeds the number of training data samples, the SVM will underperform.

**Emotion detection from infant facial expressions and cries**

The images selected for analysis are recorded from NICU at MSRTH using Wi-Fi camera and mobile camera. 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. Videos are recorded at a frame rate of 29.97 fps.

Infant pain expression (iCOPE) and NPAD dataset images are also used for analysis.

 A picture containing person, cellphone, ear

Description automatically generated A baby in a crib

Description automatically generated with medium confidence A person wearing goggles

Description automatically generated with low confidence

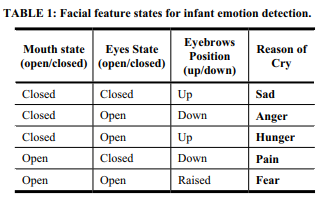
A picture containing indoor, laying, bed, baby

Description automatically generated A baby wearing sunglasses

Description automatically generated with low confidence A picture containing person

Description automatically generated A picture containing person, indoor

Description automatically generated



* The above paper uses infants images taken from internet sources and are hand labelled.
* The criteria to detect facial expressions are mentioned in the above table.
* The above criteria is checked for our dataset and it is found that most images fail to meet the above mentioned criteria.
* With the data available cry data ,classification is done into 3 classes namely hunger ,discomfort and pain.
* The scatter plot for the extraction features, namely pitch and entropy, is shown below for infant cries due to the pathological conditions hunger, discomfort and pain. The fine Gaussian kernel was found to provide
* a 79.0 % accurate classification of the given cries based on their reason for cry.
* Assuming the annotations of the cries from the chief nurse and experienced doctors from the neonatal department were accurate, the infant cries were further processed and classified. The Support Vector Machine classification used for the classification of these cries gave us an accuracy of 79%.

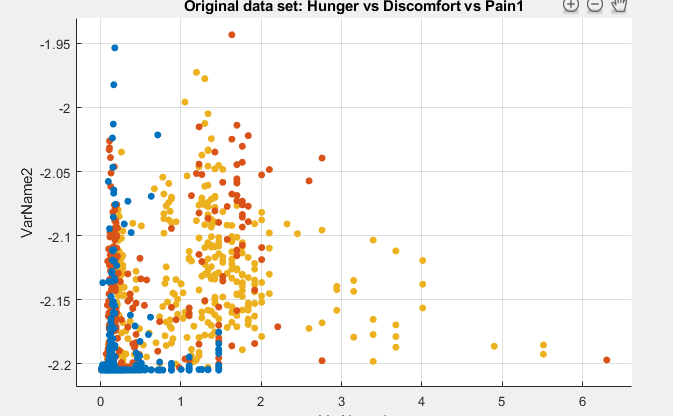
****

Figure 1Scatter Plot of SVM Classification for Hunger vs Discomfort vs Pain

****

Figure 2Confusion Matrix of SVM Classification for Hunger vs Discomfort vs Pain

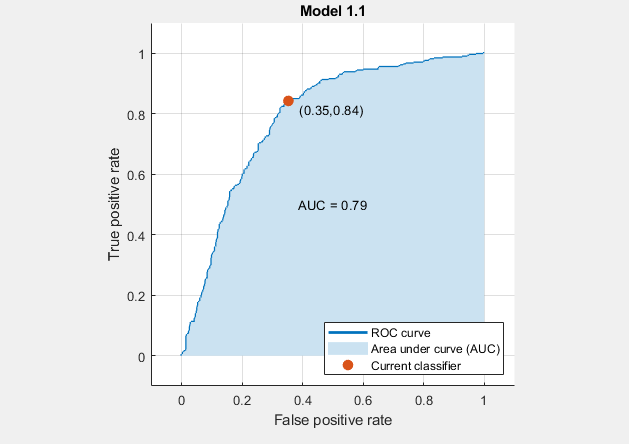


Figure 3 ROC Curve of SVM Classification for Hunger vs Discomfort vs Pain