IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received March 5, 2020, accepted March 17, 2020, date of publication March 23, 2020, date of current version April 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2982865

Novel Framework: Face Feature Selection Algorithm for Neonatal Facial and Related Attributes Recognition

MUHAMMAD AWAIS[1], CHEN CHEN[1], XI LONG[2], BIN YIN[3], ANUM NAWAZ[4], SAADULLAH FAROOQ ABBASI[1], SAEED AKBARZADEH[1], LINKAI TAO[2], CHUNMEI LU[5], LAISHUAN WANG[5], RONALD M. AARTS[2], (Fellow, IEEE), AND WEI CHEN[1,6], (Senior Member, IEEE)

[1]Center for Intelligent Medical Electronics, Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai 200433, China
[2]Department of Electrical Engineering, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands
[3]Connected Care and Personal Health Department, Philips Research, Shanghai 200433, China
[4]Shanghai Key Laboratory of Intelligent Information Processing, School of Computer Science, Fudan University, Shanghai 200433, China
[5]Department of Neonatology, Children's Hospital of Fudan University, Shanghai 200032, China
[6]Human Phenome Institute, Fudan University, Shanghai 200433, China

Corresponding authors: Chen Chen (chenchen_fd@fudan.edu.cn) and Wei Chen (w_chen@fudan.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2017YFE0112000, in part by the Shanghai Municipal Science and Technology Major Project under Grant 2017SHZDZX01, and in part by the China Postdoctoral Science Foundation Grant under Grant 2018T110346 and Grant 2018M632019.

ABSTRACT In recent times, with the advancement of digital imaging, automatic facial recognition has been intensively studied for adults, while less for neonates. Due to the miniature facial structure and facial attributes, newborn facial recognition remains a challenging area. In this paper, an automatic video-based Neonatal Face Attributes Recognition (NFAR) approach in a hierarchical framework is proposed by coalescing the intensity-based method, pose estimation, and novel dedicated neonatal Face Feature Selection (FFS) algorithm. The intensity-based method is used for face detection, followed by the facial pose estimation algorithm and FFS are dedicated to neonatal pose and face feature recognition, respectively. In this study, video-data of 19 neonates' were collected from the Children's Hospital affiliated to Fudan University, Shanghai, to evaluate the proposed NFAR approach. The results show promising performance to detect the neonatal face, pose estimation *(−45°, 45°)*, and facial features (nose, mouth, and eyes) recognition. The NFAR approach exhibits a sensitivity, accuracy, and specificity of *98.7%*, *98.5%*, and, *95.7%* respectively, for the newborn babies at the frontal (*0°*) facial region. The neonatal face and its attributes recognition can be expected to detect neonate's medical abnormalities unobtrusively by examining the variation in newborn facial texture pattern.

INDEX TERMS Neonatal face detection, facial feature selection (FFS), neonatal pose estimation, face neonatal attributes recognition (NFAR), video electroencephalogram (VEEG).

I. INTRODUCTION

Face is one of the most unique and distinct attributes of a human, which can convey relevant information such as age, gender, emotion, etc. Compared to adults, the neonatal facial structure contains approximately 10,000 nerves along with facial attributes that are still immature [1], [2]. The miniature and unformed newborn facial characteristics make it challenging to recognize their expression and sex demarcation [3]–[6]. Recently, neonate's facial rationalization has emerged as a spry area of research for various applications, e.g., baby swapping [7], baby abduction [8], neonatal pain and sedation scale via change in face pattern [9], infant pain scale measurement using facial expression moment [10], crying relating to variates in facial expression [11]–[13], etc.

However, due to the non-maturity in neonate's facial features, expressions, and random changes in their facial pattern and pose [3], [14], neonatal face and its attributes

The associate editor coordinating the review of this manuscript and approving it for publication was L. Zhang.

recognition is still a stimulating area and has only limited literature reports. Bharadwaj *et al.* [7] conducted a preliminary study for neonate's face recognition to avoid the babies swapping and abduction. Furthermore, studies on pose-invariant face recognition in newborn babies on Indian ethnics groups [8], [15] show quite promising results. Though, previous studies show that the algorithm that works well on one ethnics group may depict less accuracy on the other ethnicities as a facial pattern, demographics, and structure of neonates vary from region to region [16]. Furthermore, pose estimation and recognition of its attributes (mouth, nose, and face) have not yet been studied for neonates. Therefore, a robust algorithm is required to detect neonatal face and its characteristics precisely.

Image processing [17], [18], feature extraction [17], [19], and feature selection [20], [21] methods have been intensively studied and widely applied in face and its facial attributes detection. In the past, face detection [22], pose estimation, and facial attributes recognition were considered as an independent problem [23]. Face detection was mostly handled by the trained classifier [24]. In the last decade, numerous algorithms have been developed and claimed to have accurate performance to solve adult face detection problem, e.g., Principle Component Analysis (PCA) [25], Linear Discriminant Analysis (LDA) [26], wavelet-based algorithm [27] [28] and skin color-based algorithm (intensity-based) [29]–[32]. Among all these existing algorithm skin color-based algorithm is one of the most robust algorithm as it does not require to generate any feature matrix or Eigenface values to detect the face region. The efficacy, robustness, and device independency of the intensity-based model in the previous work motivate us to use intensity-based for neonate's face detection. On the other hand, pose detection focus on a video scenario based on 3D models [33]–[35] and facial pattern estimation was done via a classic method known as an Active Appearance Model (AMM) [36] and elastic graph matching [37]–[39]. In recent years, advancement has been made in face detection along with its pose and landmark estimation to detect a facial pose [22]. A literature review [40] shows that pose estimation models provide more precise details of facial orientation, which could be helpful in analyzing facial neonate's expression and its attributes with more accuracy. Modern research shows that face detection and pose estimation algorithms have been designed and tested for the adult face and its features recognition. The existing algorithms, e.g., intensity-based, are quite robust for face detection and pose estimation, helps to estimate the pose variation, as discussed in the previous research work. However, there are no existing algorithms that have been tested on the infant's faces and their attributes recognition using image processing algorithms to extract the face and its features from video frames.

To achieve the above goal, in this paper: a two-stage model is proposed for neonatal facial and its attributes recognition named ''Neonatal Face Attributes Recognition (NFAR)'' framework as shown in Fig. 1. At the first stage, neonates face

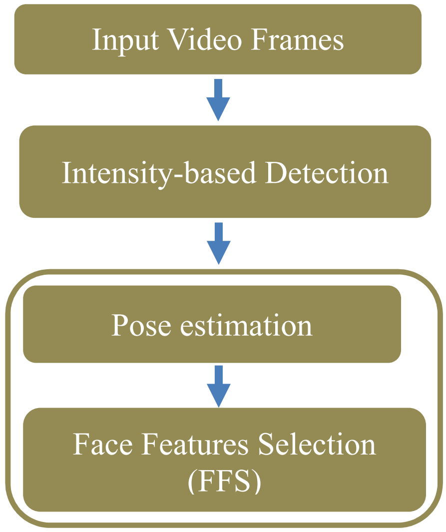


FIGURE 1. Block diagram of our proposed NFAR.

detection based on the intensity-based approach identifying the neonate's face from videos that result in removing the noisy region from each frame and pose estimation model detecting the neonate's facial pose from $-45°$ to $45°$ are performed. Next, the facial attributes recognition algorithm based on Face Feature Selection (FFS) is developed to recognize the particular Region of Interest (ROI) (e.g., eyes, mouth, and nose). One of the main advantages of our proposed hierarchical framework is that each step act as an aided tool for the following algorithm for accurate and precise detection and recognition ROI. Overall our proposed NFAR algorithm is quite robust to detect and recognize neonatal face and its attributes. In the end, the proposed framework has been compared with the existing work on neonates followed by the analysis and implementation of state-of-the-art existing algorithm on our database.

The main contributions of this research for neonatal face and its attributes recognition are summarized as below.

1) In this research, firstly, intensity-based detection and pose estimation are performed on the neonate's facial pattern independently. Secondly, we present a novel two-stage NFAR framework for face and pose estimation algorithm using video-frame imaging.

2) Dedicated neonatal pose features extraction has been designed to extract infant's faces and their attributes recognition using Face Features Selection (FFS) algorithm to achieve better performance.

3) Till now, to the best of our knowledge, there is no public newborn database available for research on the neonatal facial attributes recognition. Thus, evaluating the reliability, efficiency of neonatal facial and related attributes recognition algorithms, and providing a comparative evaluation of the performances of different algorithms are quite challenging. However, the collected dataset that with quality content can be used as a benchmark for providing a comparative comparison of the performance of the algorithm and promoting the development of the relevant algorithms.

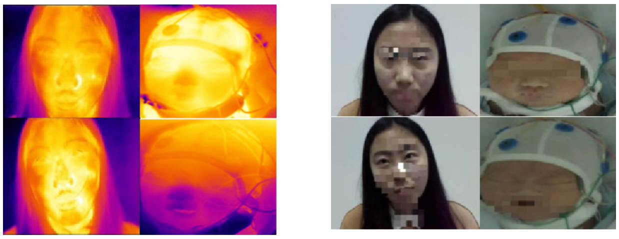


FIGURE 2. The facial characteristic of newly born babies is not proportionally equivalent to an adult face, as observed in the above illustrations.Thermal(ironbow) images shown that adult feature variety entirely differs from mature; heat variation in each feature can easily be absorbed with the naked eye.

FIGURE 3. A) Indian ethnic group [7], B) Chinese ethnic group.

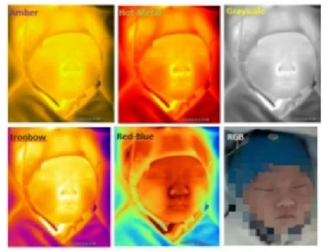


FIGURE 4. Fluke® TiX580 palette formation.

The rest of the paper is organized as follows. The next sub-sections provide background studies via a literature review on the related work and contribution of this research article. Section II describes the neonatal faces characteristic. Section III presents a detailed explanation of newborn video-database information. Section IV shows the proposed methodology to detect and recognize the neonatal face and its attributes. Section V illustrates the results of our proposed NFAR method. Section VI offers a comprehensive analysis of our proposed work with the existing algorithms. Finally, the last part concludes the paper and also provides the future direction for our research work.

II. NEONATAL FACIAL CHARACTERISTICS IN CHINESE ETHNIC GROUP

Every human face has its unique facial characters and subtle differences in form, the ratio of hard and soft facial tissues, and even topographical delineations. However, the neonate's facial structures are even physically diverse from grownup faces, and that makes it difficult to detect their facial features. The following observations and studies provide evidence to support that the existing adult faces and features detection algorithm might not work as good on the neonatal face as they do on mature adult faces. To detect the infant's face and attributes recognition through computer vision, it's vital to classify those facial features that lead to unique and discriminative features that enable detection. The face region, especially around the nasal cavity, is usually an essential point in the neonate's facial structure as the adjoining curvatures depend on it for support, as shown in Fig. 2. Furthermore, the craniofacial architecture of newly born babies has prominent eyes, small jaws, fluffy cheeks, and a soft forehead. These variations indicate that the shape and architecture of newly born babies are different from adult faces. As compared to other ethnics groups [7], the Chinese newborn baby's features are not as prominent as could easily be absorbed in different ethnic groups, as shown in Fig. 3. It has been noticed that the nasal cavity region is small and elongated fewer hairs on cheek and forehead, eyes region is small.

As compared to other ethnics groups [7], the Chinese newborn baby's features are not as prominent as could easily be absorbed in different ethnic groups, as shown in Fig. 3. It has been noticed that the nasal cavity region is small and elongated fewer hairs on cheek and forehead, eyes region is small. These observations indicated that Chinese neonatal faces also possess unique distinctive and specific facial features that could be helpful for unobtrusive, medical assistance to help clinicians to detect and diagnose neonatal abnormalities.

III. NEONATAL DATABASE

The neonatal video/images database has searched in January 2019 with the following key terms, neonatal baby's database, newly born babies images infant facial database, newborn baby's database and neonate's video database. Our neonatal video data collection was carried out at the Children's Hospital affiliated to Fudan University from November 2017 to December 2018. The experiment protocol was designed according to hospital clinical study regulation. Nineteen subjects were involved in this study. 10 out of 19 are less than 13 days old (mean = 9.8, SD = 1.87), and remaining are less than 25 days old (mean = 18.5, SD = 3.80). Video data were recorded to design the neonatal facial and its attributes detection algorithm. While collecting the database, VEEG data was recorded in parallel with videos with the future aim of our research work to predict neonatal abnormalities and behavioral monitoring by analyzing the facial motor neuron variation unobtrusively. Video data were collected from 9 am to 11-30 am for each baby, except the moments when babies had to go through the medical examination by the doctor, cluster feeding, and physical body's examination like body temperature, etc. Thus at the end of the experiment, we had 2 hours of VEEG and video data per baby.

To collect the neonate's video, Fluke®TiX580 is used [41]. One of the main advantages of using the Fluke®TiX580 camera is that it can record multiple types of color palette, as shown in Fig. 4. During the data collection, the distance between the subject and the camera is fixed; we used the Laser Pointer/Distance Finder to measure the range, from the Imager to a target, as shown

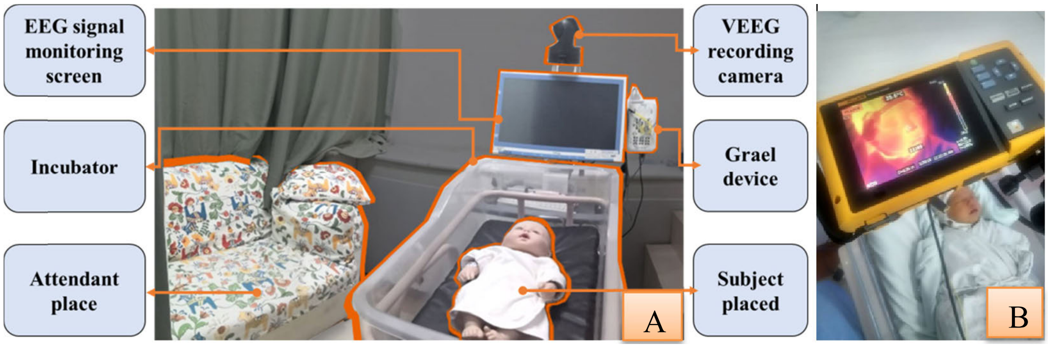


FIGURE 5. Data collection setup A) shows four different sides of the private room along with various paraphernalia, B) camera set up to record video.

FIGURE 6. A) A teat without the feeding bottle while collecting data. B) Cushion around the head to reduce random head movement.

FIGURE 7. Neonatal database challenges show the noncooperative behavior of infants that effects traditional face detection and extraction algorithms.

in Fig. 5 (B). The distance between the subject and the camera is between 0.25-0.36m. In consideration to the subject privacy, the neonatal video frames used in this paper are kept bit blurred from the specific region of the face, except in pose estimation, as it's necessary to show a whole facial area with detected points.

Newborn data collection is a bit complicated process as compared to an adult. When babies snivels, pediatrician nurse puts a teat without the feeding bottle, have been feed into the baby mouth (to make them comfortable) that results in covering the mouth/chin region. As we are using a single camera, it gets more challenging to detect the whole neonates' face using image processing algorithms, as shown in Fig. 6 (A). Another problem, while collecting neonatal data, is the baby's head movement, as they change their position left to right very quickly, that will cause a problem during facial analysis. This problem was solved by providing an extra cushion so that their head remains straight or tilt slightly left to right, as shown in Fig. 6 (B). Data was collected in a private room previously used for VEEG recording. The ward is situated and designed in such a way that it gets less impact from the hospital environment. Fig. 5 (A) shows the experimental setup for neonatal data collection. The equipment set-up details are: 1) subject position. 2) EEG signal monitoring screen. 3) VEEG recording camera. 4) Polysomnography (PSG) device, it records signals from the electrodes and sensors applied to the subject. 5): incubator. 6) Side place from where clinicians observed data quality. Camera setup, whereas collecting data have been shown in Fig. 5 (B).

During data collection, we have observed some unique challenges with infant faces that can deter traditional approaches to face recognition. The random infant movement causes Electrooculography (EOG) and electroencephalogram (EEG) electrodes to move from its position, and sometimes neonates cover the eyes and cheek region with their hands, as shown in Fig. 7. These issue has carefully monitored, video and VEEG recording have stopped, and pediatrician adjusts the position of electrodes before continue recording.

IV. METHOD: NEONATAL FACE AND ITS FEATURES DETECTION APPROACH(NFAR)

This section describes our proposed NFAR approach that involves neonatal face detection and pose estimation, followed by our proposed FFS algorithm to recognize face attributes. Fig. 8 shows the illustration of the steps involved in our proposed algorithm for face and its attributes recognition. Input video frames acts as an input for intensity-based detection to detect the neonatal facial region, then the detected face region is used by pose estimation to estimate the pose and facial area, respectively. At the end, FFS used the facial detected (pose estimated) region to recognize neonate's facial attributes.

A. INTENSITY-BASED DETECTION

The International Commission defined the intensity-based detection on Illumination in 1976 [42]. It also is known as CIE. L^*a^*b is often abbreviated as merely "Lab" color space. It articulates color as three scientific standards, L represents lightness and a and b for the green-red and blue-yellow color regions [43]. Intensity-based detection was designed to be perceptually constant concerning human facial color images.

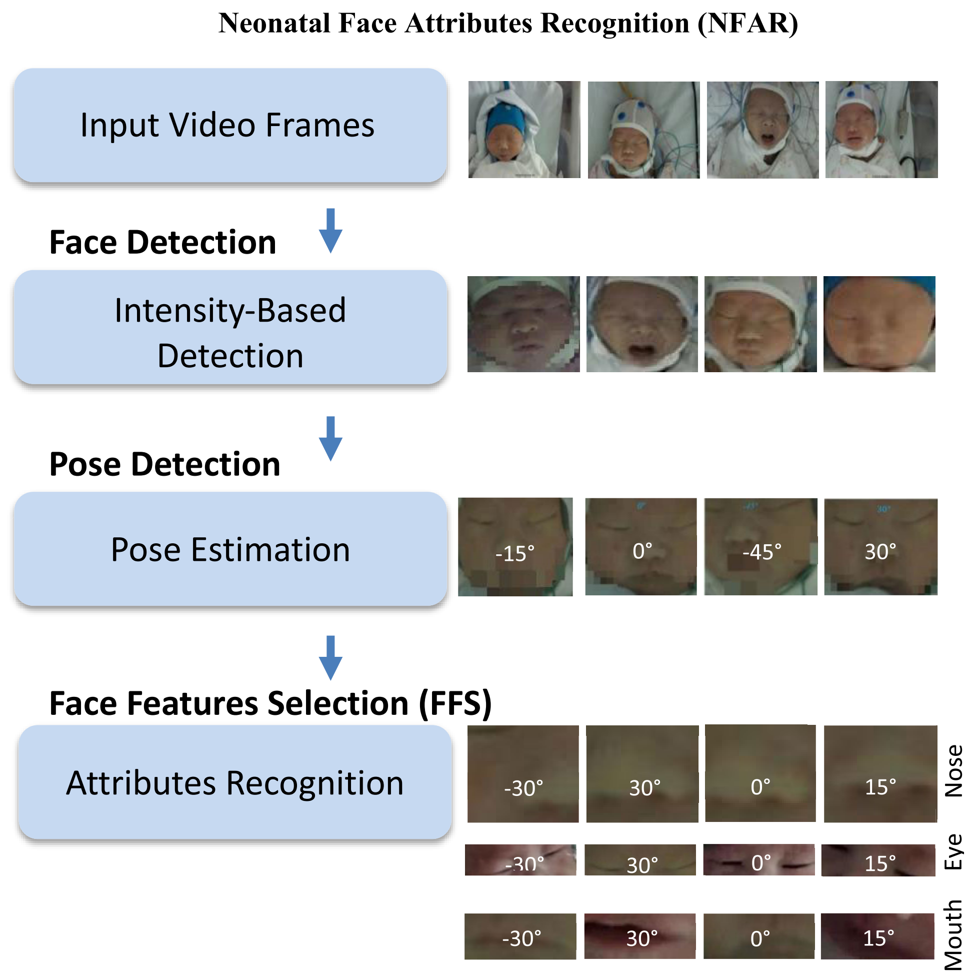


FIGURE 8. Illustrating the steps involved in the neonate's face and its attributes recognition algorithm.

One of the essential attributes of the intensity-based detection model is that it is independent of the device; it defines colors independent of how they are created or displayed. Space itself is a three-dimensional real number space that gives us space to represent an infinite number of colors. The intensity-based detection has its advantages as compared to other color models, e.g., RGB and CMYK, it aspires to perceptual consistency, and its L part closely matches the human perception of lightness. Thus, it can be used to make accurate color balance corrections by modifying output curves in the a and b components, or to adjust the lightness contrast using the L component. Intensity-based colors are defined relative to the white point of the CIEXYZ space from which they transform; thus, CIELAB values do not represent full colors unless the white area is also specified. CIELAB–CIEXYZ conversions: (Forward transformation).

$$L^* = 116f\left(\frac{Y}{Y_n}\right) - 16 \tag{1}$$

$$a^* = 500\left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \tag{2}$$

$$b^* = 500\left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \tag{3}$$

where

$$f(t) = \left\{ \begin{array}{ll} \sqrt[3]{t} & \text{if } t > \delta^3 \\ \frac{t}{3\delta^3} + \frac{4}{29} & \text{otherwise} \end{array} \right\} \tag{4}$$

Here X_n, Y_n and Z_n are the CIE-XYZ tristimulus values of the white reference point, where $\delta = \frac{6}{29}$ and t control the slope of the function (f) [40]. In our proposed algorithm, intensity-based detection takes raw image frames as input to remove the noise region and detect the neonate's facial area. As the intensity-based method is independent of the device, and due to its robustness, it is able to identify the region of interest (ROI) efficiently.

B. FACE DETECTION AND POSE ESTIMATION

Face feature selection, pose estimation, and face detection have conventionally considered as different problems with a different set of methods, such as trained classifier scanning window, view-based Eigenspace methods, and elastic graph models, respectively. For face detection and pose estimation, we have used encoding elastic deformation and three-dimensional structure, this technique involves share pools of the part with a mixture of trees. The global trees mixtures are used to model topological facial changes due to viewpoints. The pattern generated from tree mixture helps the model to analyze a large number of the facial region with low complexity [25], [28]. Overall face detection and pose estimation model is described as follows:

Model: Face and pose estimation model contains an assortment of trees that include a facial landmark shared pool of parts ($V = K$). Facial landmark has been considering part of the model, and the global mixture is used to identify the topological changes due to changes in viewpoints.

1) TREE STRUCTURE MODEL

Each tree $U_m = (V_m, E_m)$ acts as a linear- parameterized model [29], where m indicates a mixture $V_m \subseteq V$. In case of an image (I), so $l_i = (x_i, y_i)$ for the pixel location in a part of i.So scoring a ring of the configuration of part $L = \{l_i : i \in V\}$ is defined as:

$$S\,(I, L, m) = App_m\,(I, L) + Shape_m\,(L) + \alpha^m \qquad (5)$$

where

$$App_m\,(I, L) = \sum\nolimits_{i,j\in E_m} w_i^m.\varphi(I, l_i) \qquad (6)$$

Equation (6) shows the summation of appearance for placing of the template w_i^m for part i, for tuning the mixture m, at location l_i. The Histogram of Gradient (HoG) description is used as a feature vector shown in as $\varphi(I, l_i)$ extracted from pixel location l_i the image I.

$$Shape_m\,(L) = \sum\nolimits_{i,j\in E_m} a_{ij}^m.dx^2 + b_{ij}^m.dx + c_{ij}^m.dy^2 + d_{ij}^m.dy \qquad (7)$$

Equation (7) score the mixture- specific partial arrangement of parts L. where $dx = x_i - x_j$ and $dy = y_i - y_j$ are the displacement of the ith part relative to the jth part. Each term in the sum can be interpreted as a spring that introduces spatial constraints between a pair of parts, where the parameters (a; b; c; d) specify the rest location and rigidity of each spring [44] and the last term α^m is a scalar bias or ''prior'' associated with view point mixture m.

2) PART SHARING

For each mixture/viewpoints, m of part i (5) requires a separate template w_i^m. On the other hand, small changes across in viewpoint look consistent, even in extreme cases ''fully shared'' model is used as a single template for any particular change across all viewpoints $w_i^m = w$. The range between these two extremes can, as $w_i^{f(m)}$, where $f(m)$ is a function that maps a mixture index (from 1 to M) to a smaller template index (from 1 to M'). We explore various values of M': no sharing ($M' = M$), sharing across neighboring views, and sharing across all views ($M' = 1$).

3) INFERENCE

Inference corresponds to maximizing $S\,(I, L, m)$ in (5) over L and m:

$$S^*\,(I) = \overset{max}{m}\,(\overset{\max}{L}\,(S\,(I, L, m)) \qquad (8)$$

Just enumerate all mixtures, and for each combination, find the best configuration of parts. Since each mix $T_m = (V_m, E_m)$ is a tree, and the inner maximization can be done efficiently with dynamic programming.

C. NEONATAL FACE FEATURE SELECTION (FFS)

Before recognizing the face and its attributes, Euclidian (U) distance of all the points in features matrix is calculated, if the values of U is less than Discarded (D) frames, these neonatal frames are considered as False Negative (FN). The values of D may vary from as other facial datasets; it depends on camera resolution [45]. In our case, we set the value of D is 100. This step is essential to remove false positive face detection in the pose estimation step.

1) FACE EXTRACTION FROM POSE DETECTION

Once the pose (F) is detected, the first important step is to extract the face region so the image $(I)I = (x_i, y_i)$, where M, N is the number of rows and columns, respectively. F is the feature matrix that contains all the coordinates' features points found by pose detection. So $F\_S((x_i, y_i) = min(F((x_i, y_i)), F\_E((x_i, y_i) = max(F((x_i, y_i)), F\_S$ is the starting point co-ordinate for face, F_E is the endpoint of the face region. The Face is extracted by joining the row and column found in F_S, F_E in rectangle form. Once the face is detected, F is updated by discarded, those feature points lie on face boundary.

$$U\,(T \leq U) = \begin{cases} I\,(INDEX\,(M, N)) & if\ I\,(INDEX\,(M, N)) = F\,(M, N) \\ 0 & else \end{cases} \qquad (9)$$

Equation (10) helps to detect and extract the mouth/eyes region precisely by calculating the shortest Euclidian (U) of a particular index by comparing it with the value of T.

2) FFS-ALGORITHM FOR MOUTH AND EYES EXTRACTION (VOTING RULE)

For eyes region, directly used the image, on the other hand, for mouth detection horizontal flip the image to detect and extract the mouth region. Let us consider $O = M + 1$ and $P = N + 1$ according to value M, N at a particular instance in (10), to calculate the distance of specific feature point with others points so

$$(ML) = \left\{\left(Eul(U_{M,N}, U_{O,P})\right)\ if\ distance \leq (T = k)\right\} \qquad (10)$$

Once the eyes and mouth have been extracted, the rest feature matrix belongs to the nose region. The value of T has been determined from the feature region matrix (k) obtained

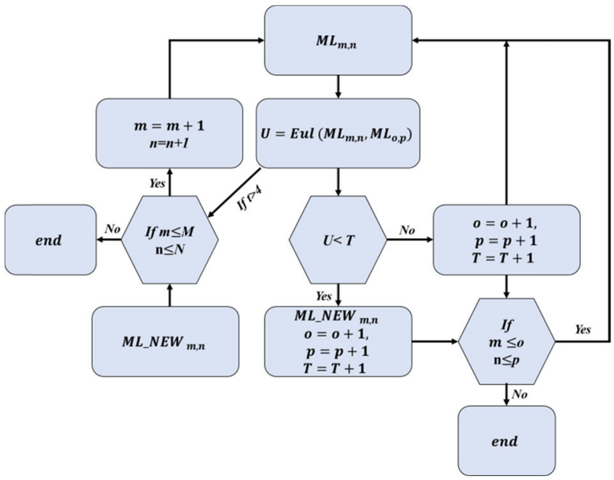


FIGURE 9. Overall flow chart of FFS.

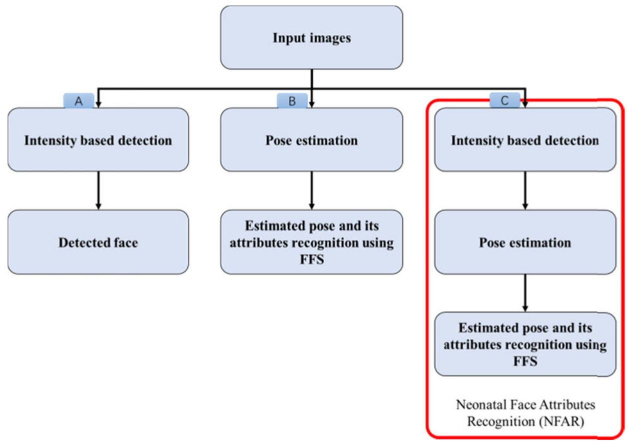


FIGURE 10. The overall description of the existing algorithm and its effect along with our proposed (NFAR) algorithm.

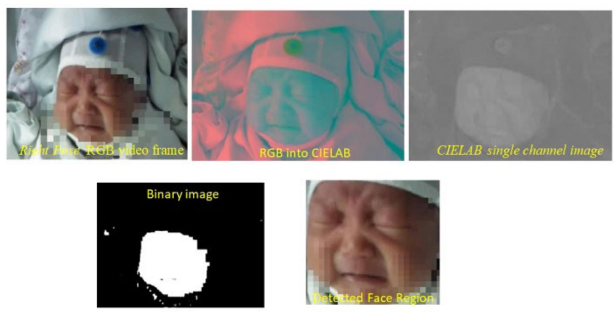


FIGURE 11. Neonates face detection using Intensity-based detection.

via pose detection. Based on fluke camera calibration and resolution, the values of k are set to 50; this value can be calibrated and adjusted according to camera resolution and pixel quality.

3) FLOWCHART FOR MOUTH, EYES AND NOSE EXTRACTION VIA FFS

Let m, n is the coordinate of the current feature point, and $O = m + 1$ and $P = n + 1$ are the next features points; first, we check these coordinates points lie inside the original image. Fig. 9 shows the overall flow chart of feature extraction where $ML_{m,n}$ is the features point matrix obtained via pose detection, $O = m+1$, and $P = n+1$ are the next features points. The U calculates the Euclidian distance between the current and the neighboring features point. When the distance is less than then T, it is considered as belong to particular face region (nose, eyes, mouth); otherwise, the value of *m, n, P, O* is incremented, and *ML_NEW* get next value from *F_S*, this process continues until the framework gets minimum four neighboring's features points from *F_S* to detect and recognize the particular facial features.

V. RESULTS

In this section, we investigate and evaluate the performance of existing face detection algorithm methods (intensity-based, pose estimation) and our proposed algorithms named "Neonatal Face Attributes Recognition (NFAR)." for the infant's face, and its attributes recognition. Fig. 10 depicts the complete description of our proposed approach adopted to evaluate the results on the neonate's dataset, where A) Intensity-based method solely detects the neonatal face region. B) Pose detection and FFS estimates the pose and recognizes facial features, respectively. C) Our proposed framework (NFAR) simultaneously detects the infant's facial region, followed by the pose estimation and FFS to estimates the pose and recognize infant's facial features precisely.

A. INTENSITY-BASED DETECTION

Our research on infant facial detection, pose estimation, and its attributes recognition rely on face detection with minimum noise region around the infant's facial area. To detect the infant's facial region using intensity-based detection, first, we have converted the RGB video frames into CIELAB color space, after analyzing the color intensity values of each CIELAB channel, the threshold has been determined. The CIELAB single-channel frame is converted into binary frames, and the connected area is separate from each other. In the end, the facial intensity region with the highest number of the related/linked part is imbricate on the original RGB and thermal images. The experiment has been performed on 700k video frames approximately from 19 infants to detect the neonatal face. We obtained Sensitivity *(Se)*, Specificity *(Sp)*, and Accuracy (*Ac*) of 99%, 99.8%, and 95.8% respectively by analyzing each frame using following Eq. (11), (12) and (13) respectively. Neonatal intensity-based detection is shown in Fig. 11, in which the RGB image is converted into CIELAB using the RGB to CIELAB transformation followed by the conversation into a binary image. The binary image is mapped to the RGB frame to detect the facial region. The results show that intensity-based detection identifies the face region with more accuracy, but it's unable to detect facial features.

$$Se = \frac{True\ Positive(TP)}{True\ Positive\ (TP) + False\ Negtive(FN)} \tag{11}$$

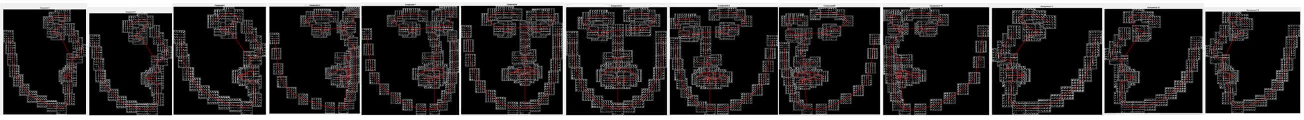


FIGURE 12. **Pose estimation: Trees model mixture encodes topological changes due to the viewpoint. Red lines signify springs between pairs of parts. All trees make use of a common, shared pool of part templates, which makes learning and inference efficient.**

$$Sp = \frac{True\ Negative(TN)}{True\ Negative(TN) + False\ positive(FP)} \tag{12}$$

$$Ac = \frac{TP + TN}{TP + FN + TN + FP} \tag{13}$$

$TP =$ correctly identified, $FP =$ incorrectly identified
$TN =$ correctly rejected, $FN =$ incorrectly rejected

Statistical results of all the face detection are calculated on the bases of the following parameter: TP = Face region exist and correctly identified by intensity-based algorithm, FP = incorrectly identified (face doesn't exist, but frames has been identified as face region), TN = neonate's face region doesn't exist and its correctly rejected by intensity-based algorithm, FN = incorrectly rejected (face exists, but frames have not been identified as face region).

B. POSE DETECTION AND FFS

Pose detection outperforms state-of-the-art detection algorithms and gives us a wide range of face detection in terms of angle on adult datasets [44]. In consideration of that, we have employed this model of pose estimation to detected face pose and followed by our dedicated neonatal algorithm known as FFS for recognizing the facial attributes. Face detection, along with the pose detection and neonatal feature recognition algorithm, improves the classification stage to have a separate analysis of face origination at $90°$ or $-90°$ as they are the counterpart of each other. The entire 19 subjects with approx. 2-hour video at each frame followed by the feature extraction has been tested using pose estimation and FFS approach.

Fig. 12 shows the standard trees model mixture generated that encodes topological changes due to the viewpoint at different angles [$-90°$, $90°$]. Red lines signify springs between pairs of parts. All trees make use of a standard, shared pool of part templates, which makes learning and inference efficient to detect and estimate the facial pose [44]. The pose estimation tree model has been used to detect the RGB face with different pose angles from [$-90°$, $90°$], as shown in Fig. 13. Once the face is detected, FFS is used to recognize the facial attributes like nose [$-30°$, $30°$], eyes [$-60°$, $60°$], and mouth [$-45°$, $45°$] as shown in Fig. 14, Fig. 15, and Fig. 16 respectively. We observed that FFS at different pose angles is getting narrow as neonates face moving toward left and right. The statistical result has been shown in Table 1. Overall, analytical results are not up to quite promising; even at specific pose angles, the facial area is not detected accurately, as shown in Fig. 13 ($-15°$, bottom left); this also results in the limited facial attributes pose detection angles. These statistical have been calculated via equations (11), (12), and (13).

FIGURE 13. **Extracted face region via FFS with different neonatal pose variations.**

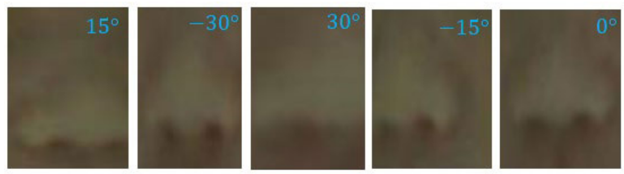


FIGURE 14. **Nose extracted region via FFS with different neonatal pose variations.**

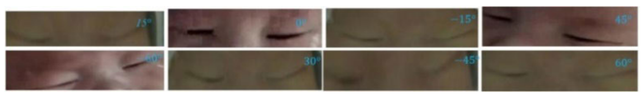


FIGURE 15. **Eyes extracted region via FFS with different neonatal pose variations.**

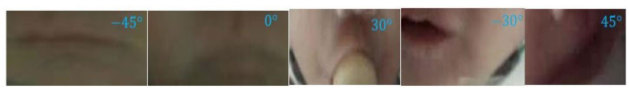


FIGURE 16. **Mouth extracted region via FFS with different neonatal pose variations.**

TABLE 1. **Statistical performance of pose estimation and FFS (F = Face, E = Eye, N = Nose, M = Mouth).**

	TP	TN	FP	FN	*Se*%	*Ac*%	*Sp* %
F	352302	196302	96653	227335	60.7	62.8	67.0
E	230525	95202	51451	146777	61.0	62.1	64.9
N	231344	93014	53639	170958	57.5	59.0	63.4
M	241095	94100	52553	136207	63.8	63.9	64.1

C. NEONATAL FACE ATTRIBUTES RECOGNITION (NFAR)

The intensity-based detection algorithm detects only the face region; on the other hand, pose estimation, and FFS extraction provides more information about face orientation and its attributes. To acquire accurate facial recognition by the better

TABLE 2. Statistical results of our proposed NFAR algorithm.

		TP	TN	FP	FN	Sensitivity (%)	*Specificity* (%)	*Accuracy* (%)
Face	*-45°*	60755	5247	295	3102	95.14	94.67	95.1
Eye		56599	280	15	4156	93.15	94.9	93.16
Nose		55489	277	18	5266	91.33	93.8	91.34
Mouth		55675	275	20	5080	91.63	93.22	91.64
Face	*-30°*	56899	6665	302	3251	94.59	95.66	94.74
Eye		53554	285	17	3345	94.13	94.37	94.13
Nose		53349	288	14	3550	93.77	95.36	93.78
Mouth		53449	291	11	3450	93.93	96.35	93.94
Face	*-15°*	66310	10110	556	4215	94.02	94.47	94.14
Eye		62110	530	26	4200	93.36	95.53	93.67
Nose		62115	534	22	4195	93.6	96.04	93.69
Mouth		62105	536	20	4225	93.65	96.4	93.65
Face	*0°*	**250180**	**20621**	**912**	**3145**	**98.7**	**95.7**	**98.5**
Eye		**241092**	**877**	**35**	**9088**	**96.3**	**96.16**	**96.3**
Nose		**242070**	**877**	**35**	**8110**	**96.4**	**96.16**	**96.7**
Mouth		**243080**	**877**	**35**	**7100**	**96.8**	**96.16**	**97.15**
Face	*15°*	55132	9565	713	3975	93.27	93.06	93.24
Eye		52277	670	43	2855	94.08	93.96	94.81
Nose		52300	685	28	2832	94.86	96.07	94.87
Mouth		52315	680	33	2817	94.89	95.37	94.89
Face	*30°*	75680	8075	611	5236	93.52	92.96	93.47
Eye		73679	570	41	2001	97.3	93.28	97.32
Nose		72502	576	35	3178	95.82	94.27	95.78
Mouth		72882	572	39	2798	96.3	94.76	95.2
Face	*45°*	62755	5142	311	2877	95.61	94.29	95.51
Eye		59654	293	18	3101	90.72	94.21	95.05
Nose		59756	287	24	2999	95.22	92.28	94.92
Mouth		59755	286	25	3000	95.21	91.96	95.2

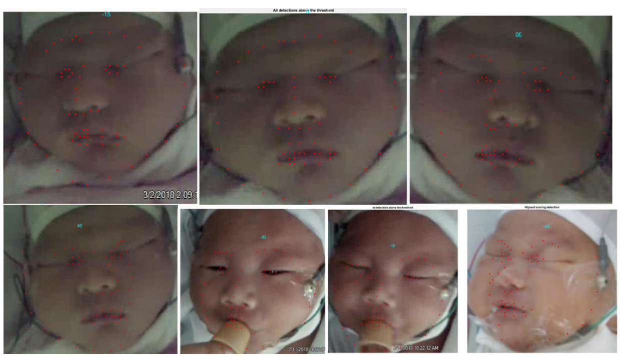


FIGURE 17. Face and its features detected region at a different angle using pose detection.

analytical result, we have designed a new framework to process these algorithms in hierarchical order. The main idea of using the NFAR approach is to reduce the noise region around the neonates' faces using the intensity-based method. So the tree mixture pose estimation model has a specific region of interest to detect the neonate's facial features precisely followed by FFS to recognize each facial feature's attributes.

The overall results are quite promising with high statistical values as compared to the raw images tested previously. Analytical results show that as face varies from left *(−45°)* to right *(+45°)* side, face poses along with its attributes have been detected more precisely. Fig. 17 shows a marked infant's face using pose estimation at a different angle. Pose estimation tree model was used to identified and FFS helps us to recognize and extract the RGB face with a varying degree of the pose, as shown in Fig. 14-16. For face attributes (mouth, nose, and eyes) recognition, it follows only those frames that are considered as TP and FP in the previous step (intensity-based face detection). Statistical results of all the attributes are calculated on the bases of the following parameter: *TP = correctly identified (face attributes exist, but features matrix (U) is more the 100), *FP = incorrectly identified (face attributes don't exist, but features matrix distance (U) is more than 100) *TN = correctly rejected ((face attributes doesn't exist, but features matrix distance (U) is less the 100), *FN = incorrectly rejected (face attributes exist, but features matrix distance (U) is less the 100). The results are manually validated by comparing it to the input frame of the particular face and its attributes. Table. 2 presents the performance of the proposed method in recognizing the neonate's face and its attributes at different angles. It achieves the best performance for detecting the face and facial attributes at the frontal (0°) facial region with the sensitivity, accuracy, and specificity of 98.7%, 98.5%, and, 95.7% respectively. As the face tilts toward the left and right, the overall accuracy decreases slightly. It is mainly because our study involves only one camera, neonates wear EEG capes and EOG electrodes for VEEG recording that cover the forehead and side region of the neonate's face. These artefacts makes it difficult

TABLE 3. Comparison of our purposed (NFAR) with existing algorithms.

Authors	Algorithm	Face detection Accuracy	Database Information	Pose information	Attributes Recognition	Ethnics Group
R. Singh and H. Om [15]	T^2WR	97.4%	Neonates	No	No	Indian
R. Singh and H. Om [49]	LBP-Gaussian level	86.9 %	Neonates	No	No	Indian
R Singh and H Om [8]	SURF	92.1 %	Neonates	No	No	Indian
Himanshu et al. [7]	SDAE	78.5%	Neonates	No	No	Indian
Tiwari et al. [47]	PCA, FLDA	80.0%	Neonates	No	No	Indian
Bharadwaj et al [48]	LBP	96.9%	Neonates	No	No	Indian
Our proposed algorithm	**NFAR**	**98.5 (0°%)**	**Neonates**	**Yes**	**Yes**	**Chinese**

TABLE 4. Comparison of the well-known existing face and its attributes detection algorithms with NFAR on our database.

	Ada-Boost (%)			*LBP (%)*			*Pose Detection (0%°)*			*Intensity Based Detection (%)*			*NFAR at (0%°) (Proposed Approach)*		
	Se	*Ac*	*Sp*	*Se*	*Ac*	*Sp*	*Se*	*Ac*	*Sp*	*Se*	*Ac*	*Sp*	*Se*	*Ac*	*Sp*
Face	35	35	34	36	36	35	60.7	62.8	67.0	99	99.8	95.8	**98.7**	**95.7**	**98.5**
Eye	No	No	No	No	No	No	61.0	62.1	64.9	No	No	No	**96.3**	**96.1**	**96.3**
nose	No	No	No	No	No	No	57.5	59.0	63.4	No	No	No	**96.4**	**96.1**	**96.7**
Mouth	No	No	No	No	No	No	63.8	63.9	64.1	No	No	No	**96.8**	**96.1**	**97.1**

for tree mixture (pose estimation) model to recognize the face and its attributes location precisely. This is also the main reason why the proposed algorithm can only detect face and facial attributes within 45° to −45°. In Table 2, slight variations can be observed at different pose angles for attributes detection between ±45°. It is because while data collection few neonates were fed with nipple (as shown in Fig. 17) by the pediatricians to comfort them, which may result in slightly less/higher identification results for one pose to another. Overall, our proposed algorithm is quite robust, with overall detection and recognition rates is less the 1 seconds approx. for single frame using Dell precision Tower 7910(Intel Xeon(R) CPU E5-2687W v4 @ 3.00GHz × 24) with Nvidia 1080Ti Graphics card.

VI. DISCUSSION

In this study, we aim to detect and recognized the neonatal facial region along with the face pose and its attributes recognition. Our proposed NFAR framework combines intensity-based, pose estimation, and FFS shows promising statistical results to recognize the neonates and its features. The advantage of using the NFAR approach acts as an aided tool to improve the recognition accuracy of the individual algorithm followed by the FFS for neonate's facial attributes recognition. Results show that when we use raw video frames directly for pose estimation and FFS, we ended up with more number of *FN* and *FP*.

Furthermore, previous research has been done on neonatal face detection [44] for different applications, e.g., to avoid babies swapping, and kidnapping [46], etc. using a hierarchical combination of various image processing algorithms. Table. 3 depicts the comparison of our proposed approach with the existing works. In the existing works, the accuracy for face detection ranges from 78.5% to 97.4%. Tiwari *et al.* [46], Bharadwaj *et al.* [47] and Singh and Om [48] conducted preliminary studies on neonates face recognition with the dedicated propose to avoid newborn swapping and abduction. Singh and Om [15] achieved the accuracy of 97.4% to detect the neonatal face. However, most of these studies can only detect face instead of recognizing pose or facial attributes. The main goal of existing works is to recognize the neonate's face and detect the discriminative features that could help them differentiate newborns from one another. In contrast, this paper provides an efficient method to discriminate the face and facial attributes effectively. Experimental results reveal that the proposed method outperforms the state-of-the-art methods.

In comparison with widely used algorithms, e.g., Local Binary Pattern (LBP) [48], Principal Component Analysis (PCA) [46], Speed Up Robust Feature (SURF) [6], Stacked Denoising Autoencoder (SDAE) [5], and pose estimation [44], we analyzed our neonatal dataset on different face detection and recognition algorithms. Table. 4 shows the performance comparison of our proposed method with the existing face recognition algorithms on the same database. The results depict that Ada-boost and LBP don't show promising results for neonatal face detection; however, intensity-based detection shows quite promising results to detect neonate's face, but it won't be able to recognize baby's facial features. Pose detection performs reasonably well to detect infant face and its features, but statistical results are still not reassuring as compared to our proposed algorithms (FAR). The main reason for these well-known algorithms doesn't perform well for infant's video frames is that the facial features are quite minute as compared to an adult where facial characteristics, e.g., eyes, lips, and nose, are pretty mature. Secondly, the existing classification model e.g., RetinaFace [49], FaceBox [50] etc., for attributes detection and recognition was trained on adult image datasets. The training of new classifiers from scratch using current algorithms with a smaller face and its attributes size could be helpful to detect the neonate's face and its features.

Although the results of the neonatal facial and its attributes recognition by our proposed approach are quite promising, the method can still be enhanced. Currently, the pose estimation model is able to estimates pose varies from $(-90°, 90°)$. However, as the face moves either toward left or right, it becomes less robust. Thus, in this paper, the pose variation is limited to $(-45°, 45°)$. Moreover, at present, the proposed algorithm was tested on 19 neonates of the Chinese ethnics group. To validate the robustness and reliability of our proposed approach, more data collection will be performed. Furthermore, current research is mainly focused on the neonatal facial and its attributes recognition. Our future aim is to analyze the facial region and its attributes recognize by our proposed method to designed an unobtrusive neonatal behavioral and abnormalities (such as epilepsy, seizure, sleep staging, etc., recognition system to ensure a more comfortable fully non-contact monitoring with no disruption at all for newly born babies by analyzing the variation in facial motor neuron using biomedical image processing. This study will be extended to unobtrusive neonatal monitoring, which can act as an aided tool to help pediatrician to predict and diagnose abnormalities, e.g., sleep disorders, seizure detection.

VII. CONCLUSION

Neonatal facial recognition is one of the complex and challenging areas of research in computer vision due to miniature facial structure and inaccessibility to newly born babies. In this paper, a novel NFAR framework for neonatal face detection and its attributes recognition by the Coalesce of the intensity-based detection, pose estimation, and dedicated novel facial attributes recognition algorithm named FFS is proposed. Intensity-based detection and pose estimation act as the face and pose detector. FFS is designed for face attributes recognition. Results exhibit that using intensity-based detection independently shows better statistical results in face detection as compared to other facial detection algorithms. However, they won't provide the pose angle and attributes recognition information. In contrast, the proposed NFAR algorithms achieve favorable results with detailed facial attributes recognition. The sensitivity, accuracy, and specificity of the proposed approach for neonate face detection can reach 98.7%, 98.5%, and 95.7% respectively from the frontal side. The accuracy of the FFS can also achieve over 96%. In the future work, the pose detection algorithm will be upgraded to detect face and its feature along with pose information with higher statistical values, especially for brisk facial movement detection. Meanwhile, instead of using a single camera, multiple cameras could be used to generate a 3D-neonatal facial pattern to analyze and absorbed minor changes in neonate's motor neurons. Moreover, at present, the proposed algorithm was tested on neonates. To validate the robustness and reliability of our proposed approach, in future, our proposed framework will be tested/validated on adult's database. Our research work has the potential for unobtrusive neonatal behavioral and abnormalities recognition to monitor events such as epilepsy, seizure, sleep staging, etc. in a way of being more comfortable, fully non-contact, with no disruption for neonate's development by analyzing facial variations.

REFERENCES

[1] L. Wenping, L. Min, H. Wenfei, and L. Yu, "A facial expression recognition method for baby video surveillance," in *Proc. 3rd Int. Conf. Multimedia Technol. (ICMT)*, 2013, pp. 516–523.
[2] S. Gupta, F. Mends, M. Hagiwara, G. Fatterpekar, and P. C. Roehm, "Imaging the facial nerve: A contemporary review," *Radiol. Res. Pract.*, vol. 2013, May 2013, Art. no. 248039.
[3] I. W. R. Bushnell, "Mother's face recognition in newborn infants: Learning and memory," *Infant Child Develop.*, vol. 10, nos. 1–2, pp. 67–74, Mar. 2001.
[4] K.Graybill. *Newborn How-to Stations Using QR Codes*. Child Development, Child Development Lessons, Family, Parenting, Technology. Accessed: Dec. 25, 2019. [Online]. Available: https://www.familyconsumersciences.com/2015/09/newborn-how-to-stations-using-qr-codes/
[5] X. Liu, W. Liang, Y. Wang, S. Li, and M. Pei, "3D head pose estimation with convolutional neural network trained on synthetic images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016.
[6] J. M. D. Barros, F. Garcia, B. Mirbach, K. Varanasi, and D. Stricker, "Combined framework for real-time head pose estimation using facial landmark detection and salient feature tracking," in *Proc. 13th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl.*, vol. 5, 2018, pp. 123–133.
[7] S. Bharadwaj, H. S. Bhatt, M. Vatsa, and R. Singh, "Domain specific learning for newborn face recognition," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1630–1641, Jul. 2016.
[8] R. Singh and H. Om, "Pose invariant face recognition for new born: Machine learning approach," in *Advances in Intelligent Systems and Computing*, vol. 1. New Delhi, India: Springer, Dec. 2015, pp. 5–6.
[9] P. Hummel, M. Puchalski, S. D. Creech, and M. G. Weiss, "Clinical reliability and validity of the N-PASS: Neonatal pain, agitation and sedation scale with prolonged pain," *J. Perinatol.*, vol. 28, no. 1, pp. 55–60, Jan. 2008.
[10] J. Lawrence, "The development of a tool to assess neonatal pain," *Neonatal Netw.*, vol. 12, no. 6, pp. 55–66, 1993.
[11] R. V. E. Grunau and K. D. Craig, "Pain expression in neonates: Facial action and cry," *Pain*, vol. 28, no. 3, pp. 395–410, Mar. 1987.
[12] S. W. Krechel and J. Bildner, "CRIES: A new neonatal postoperative pain measurement score. Initial testing of validity and reliability," *Pediatric Anesthesia*, vol. 5, no. 1, pp. 53–61, 1995.
[13] G. Zamzmi *et al.*, "A review of automated pain assessment in infants: Features, classification tasks, and databases," *IEEE Rev. Biomed. Eng.*, vol. 11, pp. 77–96, 2017.
[14] E. Kuefner, "Do all kids look alike Evidence for an other-age effect in adults," *J. Express Psychol. Hum. Percept. Perform.*, vol. 34, no. 4, pp. 811–817, 2008.
[15] R. Singh and H. Om, "(Two-dimensional)^2 whitening reconstruction for newborn recognition," *Multimedia Tool Appl.*, vol. 5, pp. 3471–3483, Aug. 2016.
[16] C. Garvie, "The perpetual line-up: Unregulated police face recognition in America," Center Privacy Technol. Georgetown Law, Washington, DC, USA, Tech. Rep. 008448, 2016.
[17] Z. Liu, J. Wang, G. Liu, and L. Zhang, "Discriminative low-rank preserving projection for dimensionality reduction," *Appl. Soft Comput.*, vol. 85, Dec. 2019, Art. no. 105768.
[18] Y. Wu, B. Jiang, and N. Lu, "A descriptor system approach for estimation of incipient faults with application to high-speed railway traction devices," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 10, pp. 2108–2118, Oct. 2019.

[19] Y. Wu, B. Jiang, and Y. Wang, ''Incipient winding fault detection and diagnosis for squirrel-cage induction motors equipped on CRH trains,'' *ISA Trans.*, to be published, doi: 10.1016/j.isatra.2019.09.020.

[20] Z. Wang, B. Du, and Y. Guo, ''Domain adaptation with neural embedding matching,'' *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Sep. 13, 2019, doi: 10.1109/TNNLS.2019.2935608.

[21] L. Zhang, Q. Zhang, B. Du, X. Huang, Y. Y. Tang, and D. Tao, ''Simultaneous spectral-spatial feature selection and extraction for hyperspectral images,'' *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 16–28, Jan. 2018.

[22] H. A. Rowley, S. Baluja, and T. Kanade, ''Neural network-based face detection,'' in *Proc. CVPR IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 20, 1996, Paper 22-38.

[23] Z. Kalal, J. G. Matas, and K. Mikolajczyk, ''Weighted sampling for large-scale boosting,'' in *Proc. Brit. Mach. Vis. Conf.*, 2008, pp. 1–8.

[24] M. Osadchy, ''Synergistic face detection and pose estimation with energy-based models,'' *J. Mach. Learn. Res.*, vol. 8, pp. 1197–1215, May 2007.

[25] L. C. Paul and A. A. Sumam, ''Face recognition using principal component analysis method,'' *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 1, no. 9, pp. 135–139, 2012.

[26] S. Elhabian and A. Farag, ''2D Face Recognition Using PCA , ICA, and LDA,'' CVIP Lab, Univ. Louisville, Electr. Comput. Eng., Louisville, KY, USA, Tech. Rep. CVIP Lab Fall-2009, 2009, pp. 1–30.

[27] L. Shen and L. Bai, ''A review on Gabor wavelets for face recognition,'' *Pattern Anal. Appl.*, vol. 9, pp. 273–292, Sep. 2006.

[28] S. Ioffe and D. Forsyth, ''Mixtures of trees for object recognition,'' in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2001, p. 2.

[29] H. M. Zangana and I. F. Al-Shaikhli, ''A new algorithm for human face detection using skin color tone,'' *IOSR J. Comput. Eng.*, vol. 11, no. 6, pp. 31–38, Jan. 2013.

[30] J. Yang, ''Skin-color modeling and adaptation,'' in *Computer Vision—ACCV*. Pittsburgh, PA, USA: Carnegie Mellon Univ., 1997.

[31] J.-C. Terrillon and S. Akamatsu, ''Comparative performance of different chrominance spaces for color segmentation and detection of human face in complex scene images,'' *Vis. Interface*, vol. 99, pp. 19–21, May 1999.

[32] C. Paper, ''Face detection algorithm with facial feature extraction for face recognition system,'' in *Proc. Int. Symp. Inf. Theory Appl.*, Sydney, NSW, Australia, Jun. 2014, 2013.

[33] W. Zhao and A. Rosenfeld, ''Face recognition: A literature survey,'' *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, 2003.

[34] V. Blanz, ''Face recognition based on a 3D morphable model,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1–12, 2003.

[35] A. Pentland, B. Moghaddam, and T. Starner, ''View-based and modular eigenspaces for face recognition,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Seattle, WA, USA, 1994, pp. 84–91.

[36] Cootes and G. Edwards, ''Active appearance models,'' *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.

[37] T. K. Leung, M. C. Burl, and P. Perona, ''Finding faces in cluttered scenes using random labeled graph matching,'' in *Proc. IEEE Int. Conf. Comput. Vis.*, Jun. 1995, pp. 637–644.

[38] L. Wiskott, ''Face recognition by elastic bunch graph matching,'' in *Proc. IEEE TPAMI*, Jul. 1997, pp. 1–23, 1999.

[39] M. Everingham, J. Sivic, and A. Zisserman, '''Hello! My name is... Buffy'—Automatic naming of characters in TV video,'' in *Proc. BMVC*, 2006, pp. 1–10.

[40] N. Ruiz, E. Chong, and J. M. Rehg, ''Fine-grained head pose estimation without keypoints,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 2074–2083.

[41] *Fluke TiX580, Expert Series Thermal Imagers*, Fluke Corp., Everett, WA, USA, 2016.

[42] N. Neophytou, ''Direct gamut editing in 3D,'' *IEEE Comput. Soc.*, May/Jun. 2008, pp. 88–98.

[43] M. D. Fairchild, *Color, and Image Appearance Models*. Hoboken, NJ, USA: Wiley, 2005.

[44] X. Zhu and D. Ramanan, ''Face detection, pose estimation, and landmark localization in the wild,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2879–2886.

[45] S. L. Tanimoto, ''Exploring mathematics with image processing,'' in *Proc. World Conf. Comput. Educ.*, vol. 6, 1995, pp. 77–85.

[46] S. Tiwari, A. Singh, and S. Kumar Singh, ''Intelligent method for face recognition of infant,'' *Int. J. Comput. Appl.*, vol. 52, no. 4, pp. 36–50, 2012.

[47] S. Bharadwaj, H. S. Bhatt, R. Singh, M. Vatsa, and S. K. Singh, ''Face recognition for newborns: A preliminary study,'' in *Proc. 4th IEEE Int. Conf. Biometrics, Theory, Appl. Syst. (BTAS)*, Sep. 2010, pp. 1–6.

[48] R. Singh and H. Om, ''Illumination invariant face recognition of newborn using single gallery image,'' *Proc. Nat. Acad. Sci., India A, Phys. Sci.*, vol. 86, no. 3, pp. 371–376, Sep. 2016.

[49] M. G. Krishna and A. Srinivasulu, ''Face detection system on AdaBoost algorithm using Haar classifiers,'' *Int. J. Mod. Eng. Res.*, vol. 2, pp. 3556–3560, Nov./Dec. 2012.

[50] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, ''RetinaFace: Single-stage dense face localisation in the wild,'' 2019, *arXiv:1905.00641*. [Online]. Available: http://arxiv.org/abs/1905.00641

MUHAMMAD AWAIS received the master's degree from Universiti Teknologi PETRONAS, Malaysia, in 2016. He is currently the Ph.D. Scholar with the Center for Intelligent Medical Electronics (CIME), Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai. His research interests include biomedical image processing, signal processing, health informatics, image processing and pattern recognition, unobtrusive-monitoring, the Internet of Things, health care, personalized, and smart environment. He received a second-year funded position in Universiti Teknologi PETRONAS, Malaysia, and the Grant Scholarship (STRIF) from the Ministry of Education, Malaysia, for excellent and efficient research, in 2015.

CHEN CHEN received the M.S. degree in the embedded system from the Institut Supérieur d'Electronique de Paris (ISEP), in 2013, and the Ph.D. degree in computer science from Université Pierre et Marie Curie (UPMC), in 2016. She is currently a Postdoctoral Researcher with the Centre of Intelligent Medical Electronics, School of Information Science and Technology, Fudan University. Her research interests include biomedical engineering, focusing on electroencephalography monitoring, biomedical signal processing, sleep analysis, wearable sensor systems, and personalized health monitoring.

XI LONG is currently an Assistant Professor with the Signal Processing Systems Group, where he advises, coordinates, and participates in many collaborative projects (with, for example, Philips, Maxima Medical Center, Kempenhaeghe, UMC Utrecht, KU Leuven, Fudan University, and so on) combining signal processing, data analysis, and healthcare. His expertise is in signal processing, time series analysis, machine learning, data analytics, deep learning, and mathematical modeling. His research interests include engineering for biomedical applications such as unobtrusive sensing, patient monitoring, vital signs monitoring, sleep, physical activity, perinatal and pregnancy monitoring, cardiorespiratory dynamics, psycho-physiological analysis, epilepsy, and brain activity. In addition, he has also published over 80 scientific articles and reports, generated more than 20 IPs or patent filings, and supervised more than 15 Ph.D. or M.Sc. students. He is a peer reviewer of more than 20 prestigious international journals/conferences in his research areas.

BIN YIN received the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China. He is currently a Principal Scientist leading the Connected Care Team with Philips Research, Shanghai. He holds a part-time professorship at the School of Information Science and Technology, Fudan University, Shanghai. After a second-year postdoctoral program with the Eindhoven University of Technology, he joined Philips Research, Eindhoven, The Netherlands. His research interests include monitoring technologies, wearables, signal processing, and intelligent algorithms, with recent applications in chronic disease management and patient engagement.

ANUM NAWAZ received the M.Sc. (Tech.) degree in information and communication science and technology from the University of Turku, Finland, and the M.Eng. degree in electronics and communication engineering from Fudan University, China, in 2018, where she is currently pursuing the Ph.D. degree with the Shanghai Key Laboratory of Intelligent Information Processing, School of Computer Science. She got a fully-funded scholarship grant by the Chinese Government (CGS) throughout her master's degree. Since 2018, she has been a Researcher with the Turku Intelligent Embedded and Robotic Systems (TIERS) Group, University of Turku. Her research interests include information security, the privacy of edge devices, blockchain, health-care, and autonomous systems.

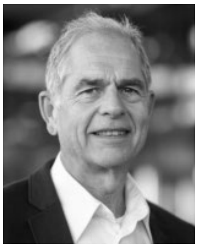
SAADULLAH FAROOQ ABBASI received the bachelor's degree from Muhammad Ali Jinnah University (MAJU), Islamabad, in 2014, and the M.S. degree in electrical engineering from the National University of Science and Technology (NUST), Islamabad, in 2017. He is currently pursuing the Ph.D. degree with the Centre of Intelligent Medical Electronics, School of Information Science and Technology, Fudan University. His research interests include biomedical engineering, focusing on electroencephalography monitoring, biomedical signal processing, sleep analysis, image processing, and neonatal health monitoring.

SAEED AKBARZADEH was born in Iran, in 1989. He received the degree in biomedical engineering from the University of Teheran Shomal, Tehran, Iran. He is currently pursuing the Ph.D. degree in biomedical engineering with the Center for Intelligence Medical Electronic(CIME), Fudan University, Shanghai, China. His scientific interests have always been concerned with the design and develop the electrical and mechanical systems, especially in biomedical engineering and neuroscience. His main research activities are particularly focused on the development and characterization of wearable devices and smart sensorized systems for biomedical engineering and the Internet of Things.

LINKAI TAO received the bachelor's degree from the Zhejiang University City College, in 2009, the master's degree from Zhejiang University, in 2012. He is currently pursuing the Ph.D. degree with the Department of Industrial Design, Eindhoven University of Technology, The Netherlands. His Ph.D. studies are focusing on the self-motivated logging system design. His research interests include user acceptance modeling, insomnia, smart healthcare, HCI, social computing technology, and motivation methods.

CHUNMEI LU received the bachelor's degree from Shanghai Jiao Tong University, in 2014. She is currently working as the Head Nurse with the Children's Hospital of Fudan University. She co-edited two books and published ten articles. Her professional fields include neonatal nursing management, neonatal vascular access management and post-discharge follow-up, and early intervention of high-risk infants. She won the Second Prize of the Science and Technology Award of the Chinese Nursing Association on Research on the Application of Key Treatment Techniques based on NIDCAP Theory in the nursing of premature infants.

LAISHUAN WANG received the M.D. degree. He was the Chief Physician, the Ph.D. Tutor, and the Deputy Director of the Department of Neonatology, Children's Hospital of Fudan University, Shanghai, China. He currently serves as the Deputy Director of the Neonatal Division of the Chinese Medical Science Branch, the Standing Committee of the Special Committee of Infant and Young Child Mental Health of the Chinese Society for Maternal and Child Health, a member of the Perinatal Critical Medicine Group of the 8th Committee of the Perinatal Medical Branch of the Chinese Medical Association, the Secretary of the Disaster Division of the Chinese Medical Science Branch, and the neonatologist of the Shanghai Paediatric Society. Currently, he undertake the National Ministry of Science and Technology key research and development program topics and the National Natural Science Foundation and other topics. He is good at the diagnosis and treatment of neonatal acute and difficult diseases and the neurodevelopmental assessment and early intervention of children with various brain injuries. He is also a Young Member and the Secretary of the Shanghai Perinatal Medical Association, Youth Editorial Board of the Chinese Journal of Contemporary Pediatrics and other magazines. He has won the Ministry of Education Science and Technology Progress Second Prize, the Chinese Medical Science and Technology Award, and the Shanghai Science and Technology Progress Second Prize.



RONALD M. AARTS (Fellow, IEEE) was born in Amsterdam, The Netherlands, in 1956. He received the B.Sc. degree in electrical engineering, in 1977, and the Ph.D. degree in physics from the Delft University of Technology, in 1995. He joined the Optics Group, Philips Research Laboratories (formerly known as the NatLab), Eindhoven, The Netherlands, in 1977, and initially investigated servos and signal processing for use in both video long play players and compact disc players. In 1984, he joined the Acoustics Group, Philips Research Labs, and worked on the development of CAD tools and signal processing for loudspeaker systems. In 1994, he became a member of the Digital Signal Processing (DSP) Group, Philips Research, and has led research projects on the improvement of sound reproduction, by exploiting DSP and psycho-acoustical phenomena. In 2003, he became a Philips Fellow at Philips Research, and extended his interests in engineering to medicine and biology in particular sensors, signal processing, and systems for ambulatory and unobtrusive-monitoring, sleep, cardiology, perinatology, drugs response monitoring (DRM), and epilepsy-detection. He has published over 400 articles and reports, and holds over 250 first patent application filings including over 175 first U.S. patent application filings and over 95 granted U.S. patents in the afore mentioned fields. Philips presented the Gilles Holst Award, in 1999, the Gold Invention Award, in 2012, the Gold with Diamond Invention Award, in 2018, and the Eureka Award, in 2001 and 2014. He has served on a number of organizing committees and as the chairman for various international conventions. He is the winner of the IEEE Chester Sall Award, in 2017; an Audio Engineering Society (AES) Silver Medal recipient, in 2010, and past governor. He has been a part-time Full Professor with the Eindhoven University of Technology (TU/e), since 2006.

WEI CHEN (Senior Member, IEEE) received the B.Eng. and M.Eng. degrees from the School of Electronics and Information Engineering, Xian Jiaotong University, China, in 1999 and 2002, respectively, and the Ph.D. degree from the Department of Electrical and Electronics Engineering, The University of Melbourne, Australia, in 2007. She worked at Bell Laboratories Germany, Alcatel-Lucent, Stuttgart, as an intern, in 2005, and she was a Research Assistant with the Department of Electrical and Electronics Engineering, The University of Melbourne, Australia, in 2007. From 2007 to 2015, she was an Assistant Professor with the Eindhoven University of Technology, The Netherlands. From 2009 to 2013, she served as the Chair of the Theme Health Care at the Department of Industrial Design, Eindhoven University of Technology. Since October 2015, she has been a Full Professor and the Director of the Center for Intelligent Medical Electronics (CIME), Department of Electronic Engineering, School of Information Science and Technology, Fudan University. Her research interests include patient health monitoring, medical monitoring system design using wearable sensors, sleep monitoring, brain activity monitoring, wireless body area networks, ambient intelligence, personalized and smart environment, smart sensor systems, and signal processing. She is an Associate Editor of the IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING and the IEEE JOURNAL ON BIOMEDICAL HEALTH INFORMATICS and the Managing Editor of *IEEE Reviews in Biomedical Engineering*.