SEMARAK ILMU PUBLISHING

Semarak Engineering Journal

Journal homepage:
https://semarakilmu.my/index.php/sej/index
ISSN: 3036-0145



Enhancing Infant Pain Cue Recognition Using a Combined Preprocessing Method under Varying Noise Conditions

Muhammad Naufal Mansor[1,*], Wan Azani Mustafa[1], Norhayati Mohd Zain[2], Moh Khairudin[3]

[1] Faculty of Electrical Engineering and Technology, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia
[2] Medical Imaging Department, Faculty of Integrative Medicine & Life Sciences, KPJ Healthcare University, 17010, Nilai, Malaysia
[3] Department of Electrical Engineering, Universitas Negeri Yogyakarta, Karangmalang, 55281 Yogyakarta, Indonesia

ARTICLE INFO

Article history:
Received 12 August 2026
Received in revised form 5 September 2026
Accepted 10 October 2026
Available online 26 October 2026

Keywords:
Infant Pain; Preprocessing; Noise

ABSTRACT

Pain can be conveyed through verbal communication or patient self-report. Nonetheless, assessing pain in infants and patients who cannot communicate verbally (such as the mute, mentally challenged, or those requiring assisted breathing) can be challenging. Despite these circumstances, only an experienced medical examiner may comprehend their non-verbal signs of pain. Moreover, pain can be troublesome at varying levels of noise. At times, even an experienced pediatrician may make an error in their assessment or overlook a minor change in an infant's face. Even when they are present, it appears challenging to recognize the subtle anxiety in the facial twitch. To address this obstacle, an automated method for assessing infant pain through facial-cue analysis is proposed. The T-test results validated the relationships among the new combinations generated by this algorithm. Ultimately, a feature fusion approach combining Principal Component Analysis (PCA), Discrete Cosine Transform (DCT) coefficients, and Local Binary Patterns (LBP) was employed. The effectiveness of the complete features descriptor and the suggested preprocessing technique was evaluated against other well-known preprocessing methods using Support Vector Machine (SVM) classifiers at various noise levels.

1. Introduction

Pain is a feeling that has both emotional and sensory components. According to the definition provided by the International Association for the Study of Pain, pain is a sensation produced in actual or potential tissue damage [1]. Expressions of pain are made through vocalizing. This may be difficult, especially in babies and patients that cannot speak out. It is difficult to express pain for the mute, mentally handicapped, or patients that require breathing support.

However, even in such cases, their muted language can be comprehensible only to an experienced medical professional. Moreover, pain-related facial cues may be difficult to recognize

* *Corresponding author.*
E-mail address: naufal@unimap.edu.my

https://doi.org/10.37934/sej.16.1.172196

under varying viewing angles, distances, and illumination levels. Even the most experienced pediatricians sometimes make mistakes or do not notice the slightest alteration on a baby's face. Although all these factors are present, it seems difficult to understand pain-related facial movements. For this reason, a valuable way to detect the pain of a newborn should be insensitive to image variations.

In particular, some of the highly advanced methods have been used in scientific research to overcome difficulties encountered while recognizing physical distress [2-6]. One of the most popular methods involves the recognition system of the pain of infants, created by S. Brahnam et al., that is based on the Classification of Pain Expressions (COPE) dataset. They recommended the usage of novel texture analysis techniques, namely, Elongated Binary Pattern (ELBP) and Elongated Ternary Pattern (ELTP), which serve as an accurate basis for pain stage classification [7]. Moreover, some of the studies, such as [8] and [9], have provided examples of using mixed AI models, including simultaneous optimization of Neural Network Simultaneous Optimization Algorithm (NNSOA) with Linear Support Vector Machine (LSVM). It is important to note that the results of their work showed that the combination of LSVM and NNSOA could provide better classification accuracy. K. Pun and Y. Moon also agreed that multifaceted classification methods provide more data about the structure of the data set [10].

Despite the wide use of such scales as NIPS, CRIES, PIPP, and N-PASS for infant pain assessment [11-17], their practical value is hampered by the presence of systemic and knowledge-related threats [11-13]. The latter include low assessment skills, absence of protocols, lack of engagement among health-care workers, and lack of knowledge about neonatal pain signs. In addition, exaggerated fears regarding the development of medication dependence and fatigue among providers hinder effective pain management [13]. A systematic review of scientific literature demonstrates that insufficient consideration was devoted to image preprocessing for recognizing infant pain. Previous research revealed the influence of changes in lighting conditions, viewing angles, distance, and image noise on face detection precision [16, 18-21]. Based on the above, the use of appropriate techniques for image processing might greatly increase the performance of facial recognition systems allowing even basic face features extraction techniques and classification algorithms to provide highly accurate results [18-20]. In addition, current approaches for infant pain detection are prone to subjective human observation and inconsistencies in procedures.

Thus, in order to resolve these issues, this paper suggests a reliable pre-processing framework for infant pain detection in various noise environments. The framework involves histogram remapping, Adaptive Single Scale Retinex (ASR), gamma correction, reflectance estimation, and contrast equalization [21,22–26]. Histogram remapping ensures image consistency due to the normalization of the intensity distribution, whereas ASR minimizes illumination effects and halo artifacts without damaging facial details. The post-processing steps further increase the quality of the images, generating more discriminating facial features.

The suggested framework was tested on the Infant COPE database and compared with a number of other well-known illumination compensation techniques. The illumination is especially crucial since degraded or inconsistently illuminated facial images can change the appearance of pain-related facial cues before classification. Therefore, a classifier can get features containing undesirable variations instead of real facial information. Thus, the problem of the research discussed in this paper is the development of an effective image processing and feature representation technique for discriminating infant facial pain cues under variable illumination and noise conditions.

The remainder of this paper is organized as follows: Section 2 describes the proposed framework, Sections 3 and 4 present the feature extraction methods and the Support Vector Machine (SVM) classifier, Section 5 discusses the experimental results, and Section 6 concludes the paper.

2. Proposed Study

As shown in Fig. 1, the proposed scheme is made up of three phases, namely, preprocessing, feature extraction, and classification. The input images are reduced to a size of 100 x 120 pixels in order to keep computation costs low. The phase of preprocessing makes use of the proposed enhancement technique in order to enhance images depending on illumination and noise. The features of faces are extracted with the help of PCA, ICA, DCT, and LBP and classified using SVM.

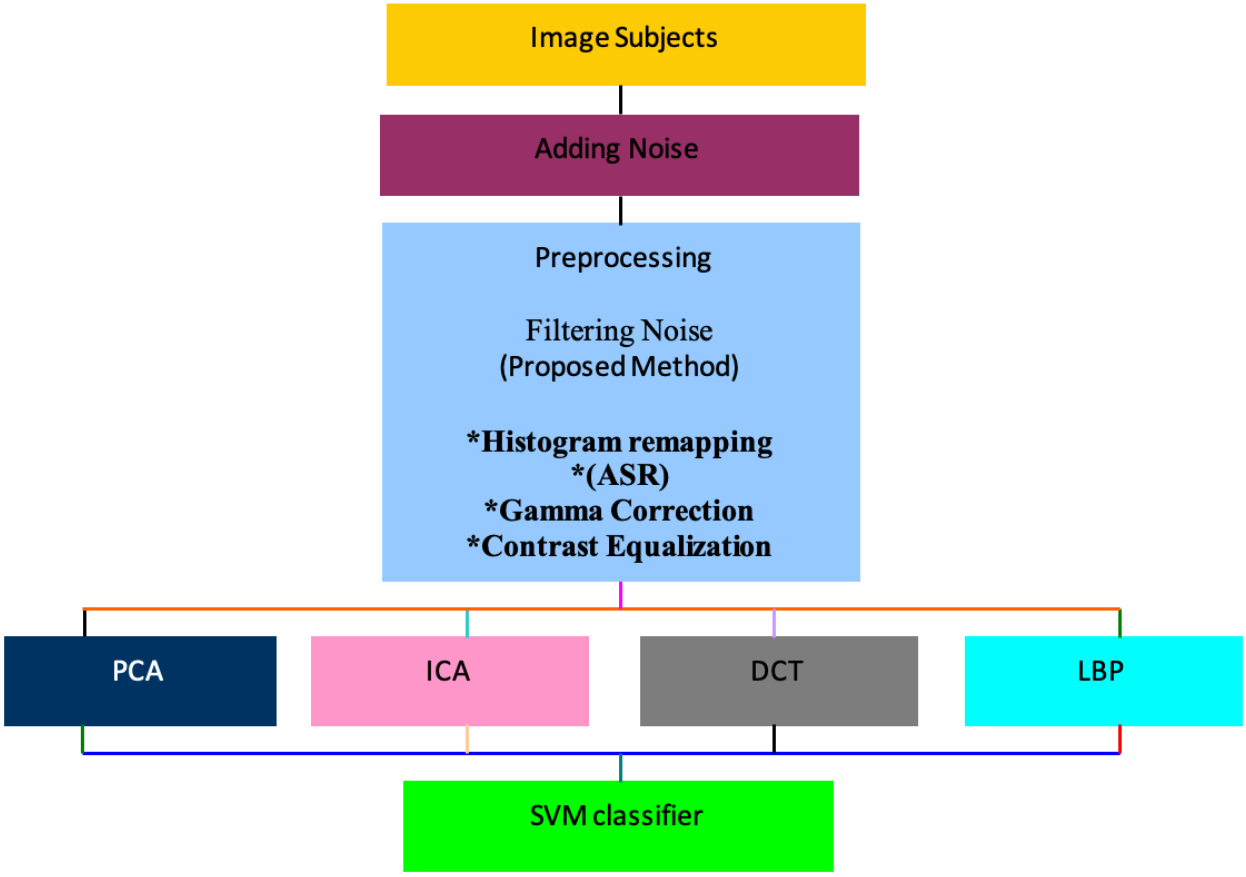


Fig. 1. The proposed Study Procedure

2.1 Image Subjects

These experiments were carried out based on the Infant COPE database, which consists of 204 facial images of 26 newborn infants (13 male and 13 female), aged between 18 hours and 3 days. The images consist of 18 cry stimuli, 67 rest stimuli, 36 friction stimuli, 23 air stimulus, and 60 pain stimuli images [2]. Data collection is made following the ethical guidelines of St. John's Health System Inc. [2]. Some example images taken from the Infant COPE database are shown in Fig. 2, and more details about the dataset and data acquisition can be found in [4], [5].



Fig. 2. Examples of the facial expressions in the Infant COPE database

2.2 Adding Noise

The most important contribution of this paper is a pre-processing technique to deal with images having high-density salt-and-pepper noise corruption. This technique can also be used as an image improvement technique in general. Salt-and-pepper noise is generated similar to [27-29]. In contrast to other techniques, the noise density is varied between 1% and 9% in steps of 2%. Fig. 3 shows the restoration of infant pain images contaminated with varying amounts of salt-and-pepper noise.



(a) (0.01 noise level) (b) (0.03 noise level) (c) (0.05 noise level) (d) (0.07 noise level) (e)(0.09 noise level)

Fig. 3. An example of the facial expressions in the Infant COPE database with different noise densities that vary from Figure 3(a) at a noise level of 0.01 to Figure 3(e) at a noise level of 0.09

2.3 Filtering

Median filters are extensively used in image processing, signal processing, and time-series processing as a means of smoothing. Unlike linear filters that are extremely sensitive to outliers and experience significant performance deterioration due to the presence of small amounts of noise, The median filter can remove high-amplitude noise in the input data. [27, 30, 31].

2.4 Histogram remapping

One of the most commonly used methods for histogram normalization is the normal or Gaussian distribution. Normal curve was first developed by De Moivre in 1733, but was later used by Laplace in 1783 for modeling errors, and Gauss in 1809 for the analysis of astronomical observations. The formula of the probability density function of the Gaussian distribution is:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \quad (1)$$

where μ is the mean, and σ is the standard deviation. In this experiment (Fig. 4), the standard normal distribution with $\mu = 0$ and $\sigma^2 = 1$ was used, as suggested by [22].

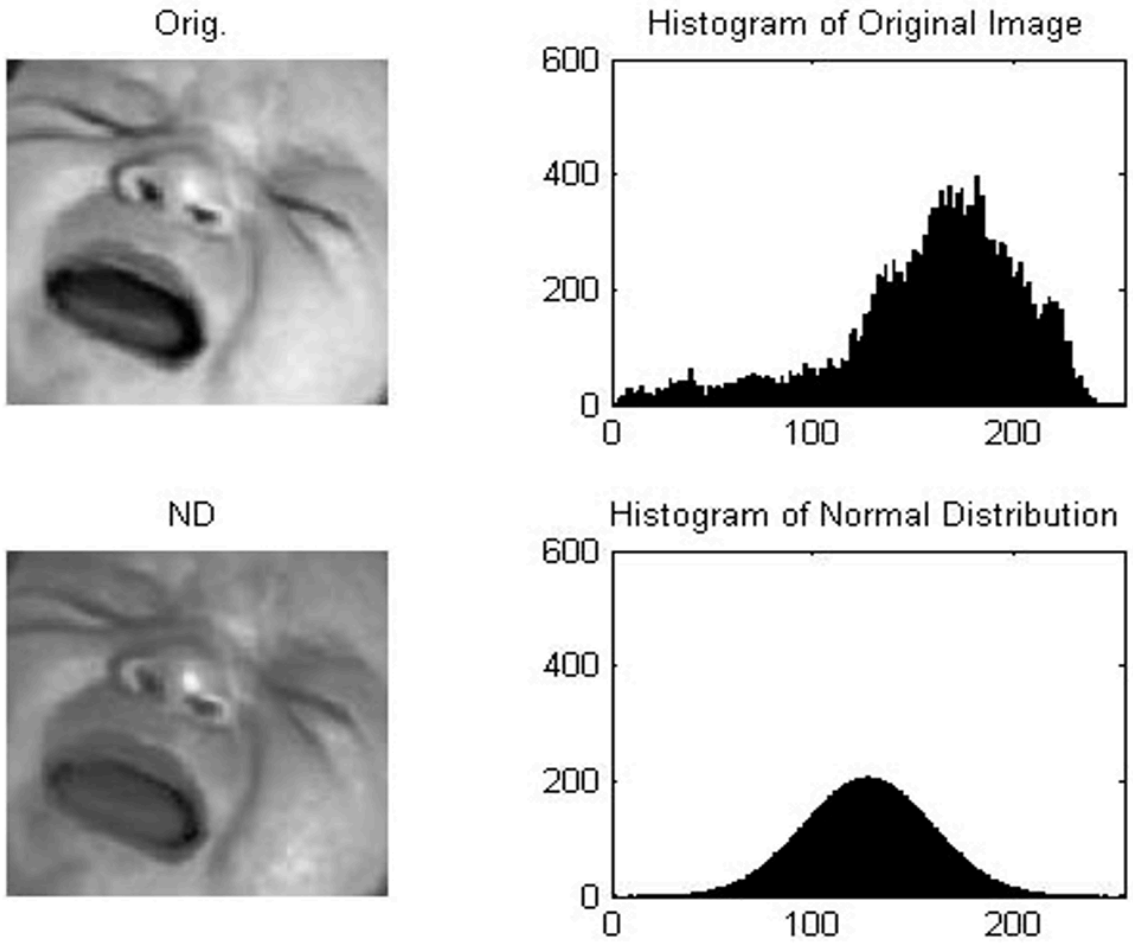


Fig. 4. A sample image and its histogram before (upper row) and after (lower row) mapping the histogram to a normal distribution

2.5 Adaptive Single Scale Retinex (ASR)

The ASR method was introduced by Parkin [14]. The method for estimating the illumination uses repeated convolution of the input image with a 3×3 smooth mask, using the coefficient and combination of two illumination values per pixel as in Equation (2).

$$R'_{ASR}(x,y) = log(I(x,y)+1) - log(L^{(T)}(x,y)+1), \quad (2)$$

where *I* denotes the input grayscale image. *T* refers to the number of iterations carried out. *T* = 15 was adopted as recommended by [14]. *R* is the reflectance and luminance *L* function (in the log domain). Figure 5 shows some ASR algorithms applied to the Infant COPE database.



Fig. 5. An example of the facial expressions in the Infant COPE database with the ASR Method

2.6 Gamma correction

Gamma correction is a nonlinear gray level mapping where new gray levels are mapped from the old gray levels. The connection between them is that the output is directly proportional to the input raised to the power of gamma. The equation used to calculate the output is:

$$I' = 255 \times \left(\frac{I}{255}\right)^{\frac{1}{\gamma}}, \quad (3)$$

where I' represents the intensity of the gamma image and γ represents the gamma. Gamma correction effect on images can be seen through the use of the original image taken with different gamma values such that if γ > 1 makes the shadow darker and γ < 1 makes the dark parts brighter. The best choice for γ is in the range of [0,0.5]. Here γ = 0.2 is chosen as a default value proposed by [21]. Some examples of Gamma Correction method are shown in Fig. 6.

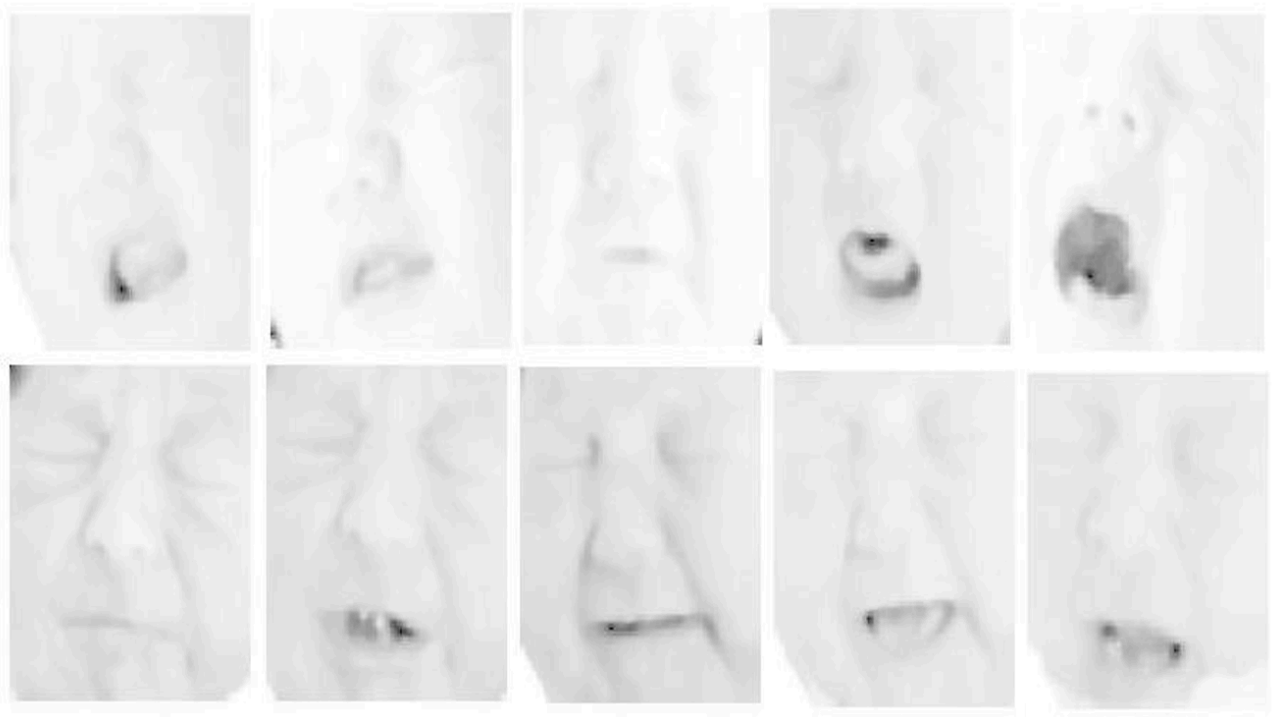


Fig. 6. An example of the facial expressions in the Infant COPE database with Gamma Correction

2.7 Contrast Equalization

Lastly, contrast equalization was performed. It is assumed that the signal generally has very high values due to highlights, shadows, and low-value regions like nostrils and cheeks. A fast and straightforward approximation using a two-step approach is given below [21]:

$$I(x,y) \leftarrow \frac{I(x,y)}{(mean(|I(x\prime,y\prime)|^{\alpha})^{\frac{1}{\alpha}}}, \tag{4}$$

$$I(x,y) \leftarrow \frac{I(x,y)}{(mean(min(\tau,|I(x\prime,y\prime)|^{\alpha}))^{\frac{1}{\alpha}}}. \tag{5}$$

Here, α is a compressive exponent. τ is a threshold to truncate the first normalization. α = 0.1 and τ = 10 were used as in [21]. Some aspects of the contrast equalization process are illustrated in Fig. 7.

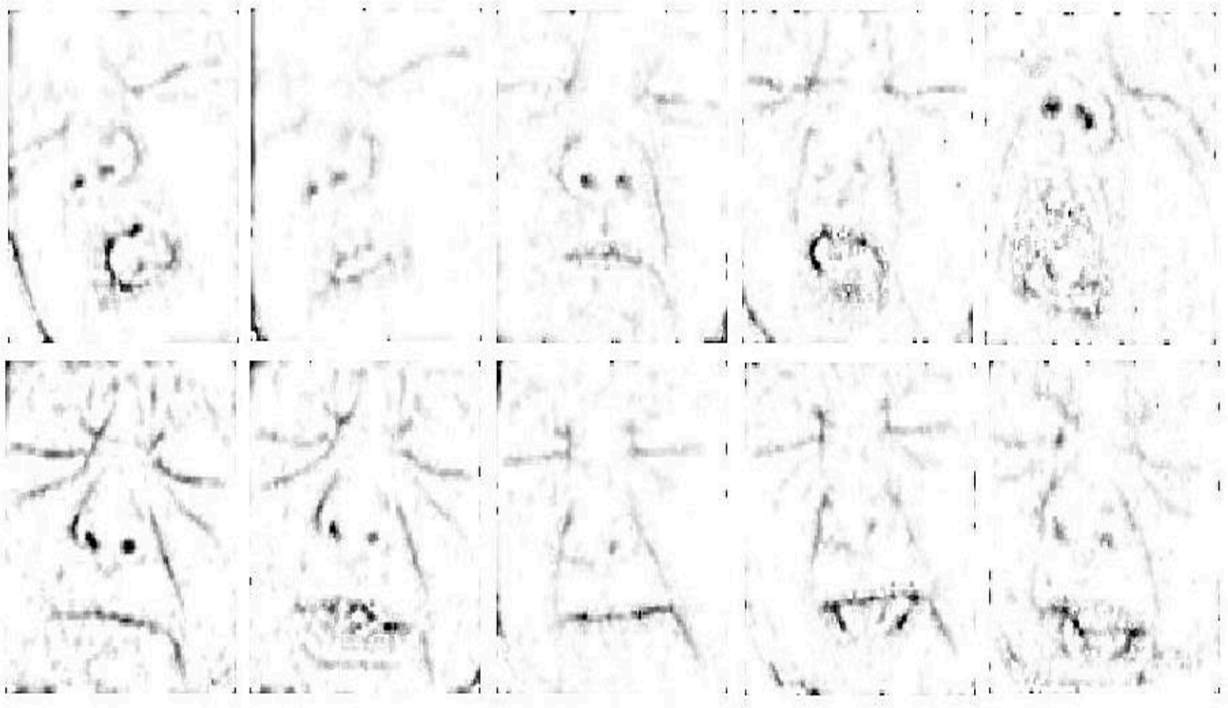
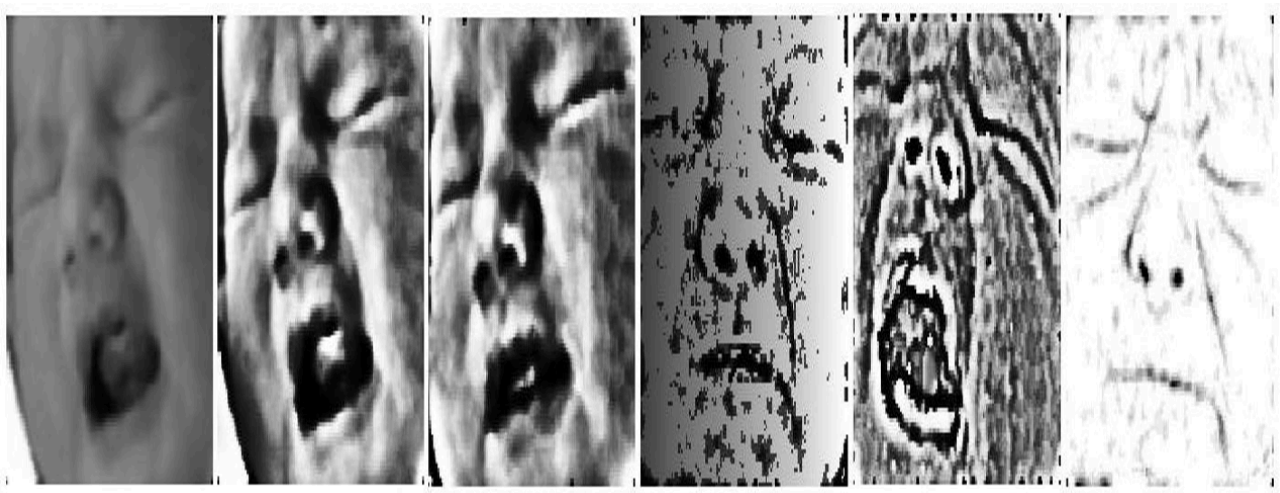


Fig. 7. An example of the facial expressions in the Infant COPE database with the proposed method

2.8 Compared algorithms

Four illumination normalization algorithms and one (None) no-preprocessing image were compared with the researcher's method. The entire method is listed below:
(None) no preprocessing
The Single-Scale Retinex (SSR) algorithm [18]
The Multiscale Retinex (MSR) algorithm [31]
The Adaptive Single-Scale Retinex (ASR) algorithm [23]
The Tan and Triggs (TT) normalization technique [21]

For comparison, some example images obtained after applying the above illumination compensation methods are presented in Fig. 8.



(NONE) (SSR) (MSR) (ASR) (TT) (OPM)

Fig. 8. Examples of the effects of the different preprocessing methods. From left to right: (None) no preprocessing; (SSR) Single Scale Retinex; (MSR) Multiscale Retinex; (ASR) Adaptive Single Scale Retinex; (TT) Tan and Triggs normalization technique, and (OPM) Our Preprocessing Method

2.9 Method Validation

A statistical analysis was conducted to evaluate the proposed OPM preprocessing method against five alternative approaches: no preprocessing, Single Scale Retinex (SSR), Multiscale Retinex (MSR), Adaptive Single Scale Retinex (ASR), and Tan and Triggs normalization. The evaluation was based on the reported t-test p-values and confidence intervals, using a significance level of 0.05.

The proposed OPM method obtained a p-value of 0.0094 and a reported confidence interval of [0.0230, 0.1630]. The p-value was below the conventional significance threshold, while the confidence interval excluded zero. These findings provide evidence of a positive difference relative to the reference condition used in the reported statistical test. Nevertheless, the magnitude and practical importance of this difference depend on the outcome variable and its measurement scale.

Among the existing approaches, MSR and SSR reported p-values of 0.0336 and 0.0435, respectively. Both methods met the unadjusted 5% significance criterion. MSR had a confidence interval of [0.1984, 0.3136], whereas SSR had an interval of [0.0025, 0.1644]. ASR did not reach statistical significance (p = 0.0709; 95% ci [−0.0482, 0.0020]). Similarly, Tan and Triggs normalization (p = 0.7699; 95% ci [−0.0823, 0.0610]) and the unprocessed images (p = 0.9997; 95% ci [−27.4761, 2.4136]) did not demonstrate statistically significant differences under their respective tests.

Overall, the findings provide preliminary statistical evidence in favour of investigating OPM as a preprocessing method for infant facial image analysis. Further validation using direct feature extraction method comparisons, and classification performance metrics will be discussed in next section to conclude that OPM is superior to the existing methods.

Table 1
t-test Analysis for various Techniques

Method	Hypothesis	*p*	*ci*
(None) no preprocessing	1	0.9997	-27.4761 2.4136
(SSR) Single Scale Retinex	1	0.0435	0.0025 0.1644
(MSR) Multiscale Retinex normalization technique	1	0.0336	0.1984 0.3136

(ASR) Adaptive Single Scale Retinex normalization technique	1	0.0709	-0.0482 0.0020
Tan and Triggs normalization technique	1	0.7699	-0.0823 0.0610
(OPM) Our preprocessing normalization technique	1	0.0094	0.0230 0.1630

3. Feature Extraction

3.1 Principal Component Analysis (PCA)

PCA [32,37] tries to find the reduced dimensionality representation of the data such that it captures as much of the data variability as possible. The analysis has been carried out by finding a linear basis of reduced dimensionality of the data and maximizing the amount of variance captured. Mathematically speaking, PCA tries to find a linear map T that maximizes $T^T\, cov_{X-\bar{X}}\, T$, where $cov_{X-\bar{X}}\, T$is the covariance matrix of the zero-mean data X. It can be seen that this linear map is constructed using the d principal eigenvectors (i.e., principal components) of the covariance matrix of the zero-mean data. Thus, PCA solves the eigenvalue problem:

$$cov\, v = \lambda v \quad (6)$$
$$X - \bar{X}$$

The eigenproblem is solved for the d principal eigenvalues λ. The corresponding eigenvectors form the columns of the linear transformation matrix, T. Moreover, the low-dimensional data representations yi of the data points xi are computed by mapping them onto the linear basis T i.e., $Y = (X - \bar{X})T$ [33]. PCA has been successfully applied across a wide range of domains, including infant pain detection [6, 38, 2, 3], face recognition [34], coin classification [35], and seismic series analysis [36, 39]. Accordingly, 70 PCA components were employed as inputs, following [6].

3.2 Discrete Cosine Transform coefficients (DCT)

The DCT or DCT-II is commonly used in signal and image processing for lossy data compression. This is because it has an energy-compaction property. Most of the signal information ends up packed into a few low-frequency DCT components. This behavior is similar to the Karhunen-Loève transform, which provides optimal decorrelation. The DCT works well for signals that come from certain Markov processes.

The two-dimensional DCT-II of blocks with size 8x8 is computed, and the output is then quantized and entropy-coded. Here, the value of n is often eight, while the formula for DCT-II is applied to rows and columns of the block. The output of this transformation is an array of 8x8 transform coefficients, where the (0,0) element represents the DC (zero-frequency) component, and other elements represent the higher vertical and horizontal spatial frequency [36].

The DCT is one of many transforms that take their input and produce a linear combination of weighted basis functions. These basis functions are commonly the frequency. The 2-D DCT is just a one-dimensional DCT applied twice, once in the x direction and again in the y direction. One can imagine the computational complexity of doing so for a large image. Thus, many algorithms, such as the Fast Fourier Transform (FFT), have been created to speed the computation. The DCT equation (Equation 7) computes the *i*, *j*th entry of the DCT of an image [41].

$$D(i,j) = \frac{1}{\sqrt{2N}} C(i)C(j) \sum_{x=0}^{N=1} \sum_{y=0}^{N=1} p(x,y) \cos\left[\frac{(2x+1)i\pi}{2N}\right] \cos\left[\frac{(2y+1)j\pi}{2N}\right], \quad (7)$$

$$C(u) = \begin{Bmatrix} \frac{1}{\sqrt{2}} & if(u) = 0 \\ 1 & if(u) > 0 \end{Bmatrix}, \quad (8)$$

where p (*x*, *y*) is the *x*, *y*th element of the image represented by the matrix *p*. *N* is the size of the block on which the DCT is done. The equation calculates one entry (*i*, *j*th) of the transformed image from the pixel values of the original image matrix [42]. Following [6], 80 DCT coefficients were used for classification.

3.3 Local Binary Patterns (LBP)

The LBP features were initially designed for texture analysis, but later came into use for analyzing facial images [7, 39, 43]. The most important characteristics of the LBP features are the simple calculations involved and the robustness to variations in illumination. LBP features can be calculated in a few seconds in one pass over the raw image and are in a low-dimensional feature space, but still provide discriminating features for facial analysis.

LBP was first proposed by Ojala et al. [44]. The algorithm computes the label of each pixel using thresholding of the 3x3 window around the pixel using the center value, providing a binary output (see Figure 9 below for illustration). It also calculates a histogram called the LBP histogram. All these histograms concatenated together provide a texture descriptor.

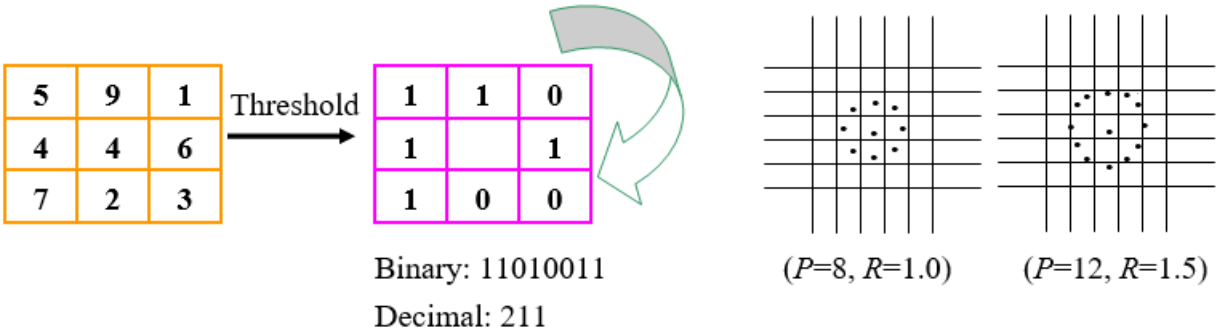


Fig. 9. Left: The basic LBP operator [46]. Right: Two examples of the extended LBP [31]: a circular (8,1) neighborhood, and a circular (12, 1.5) neighborhood

There is a problem with this simple LBP operator since, because of having a small neighborhood, it does not take the main structures into account that have a large scale. This operator was improved and can be applied to neighborhoods of various scales [45]. Therefore, it can employ a circular neighborhood. Some examples of the enhanced LBP are presented on the right side of Fig. 9 where (P, R) means P sample points of radius R.

LBP goes through a number of extensions, each one of which solves an issue in the preceding extension. Later on, a universal pattern of LBP was presented [46]. LBP is referred to as a uniform pattern, if it possesses at most two bit transitions between 0 to 1 or vice versa in the binary sequence taking into consideration its circularity. Some examples of a uniform pattern include 00000000, 001110000, and 11100001. Uniform patterns are found to comprise approximately 90% of all patterns in the (8, 1) neighborhood and 70% in the (16, 2) neighborhood of the texture image [47].

The uniform pattern is described as follows. Uniform LBP is denoted as $LBP^{u2}_{P,R}$. The (*P*, *R*) represent the sampling points and radius used for neighborhood, while (u2) indicates that only

uniform patterns are used. Other patterns are labeled as a single pattern. A histogram of a labeled image f_l (x, y) can be defined as:

$$\boldsymbol{H_i = \sum_{x,y} I(f_l(x,y) = i), i = 0, \ldots\ldots n-1,} \quad (9)$$

where n is the number of different labels produced by the LBP operator.

$$I(A) = \begin{cases} 1 & A \text{ is true} \\ 0 & A \text{ is false} \end{cases}. \quad (10)$$

The histogram defined above contains information about the distribution of edges, flat areas and spots across the whole image. Nevertheless, to create an excellent face representation, spatial information must also be collected as a feature. Thus, the face image is divided into m small regions R_0, R_1,..., R_m and a spatially enhanced histogram is defined as :

$$H_{i,j} = \sum_{x,y} I(f_l(x,y) = i)\ I((x,y) \in R_j), \quad (11)$$

where I = 0,..... n-1, j = 0,.....,m-1

As a result of the above operation, the histogram produced contains information at three levels, which are pixel-level, regional-level, and global-level. This results from concatenating all pixel-level patterns to produce a regional-level pattern. Additionally, all regional-level patterns are summed to produce a global description of the image [48].

3.4 Data Features Validation

Scatter plots of all features are generated to examine correlations at different levels of noise. Meanwhile, these scatter plots demonstrate the trend of the generated features with respect to exposure noise levels (Figs. 10-12).

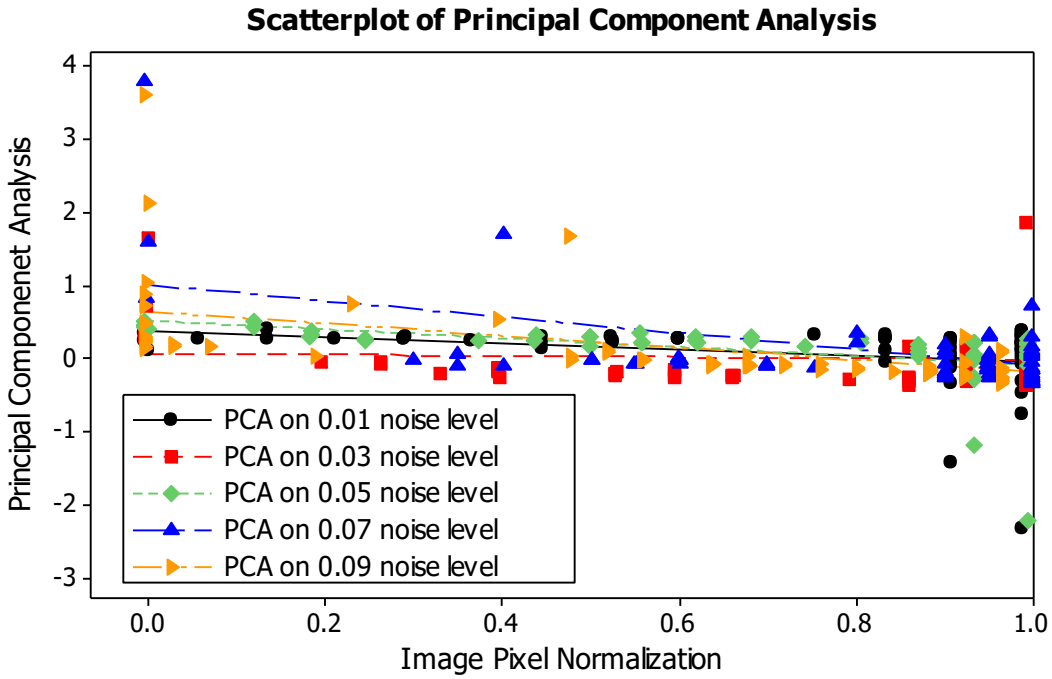


Fig. 10. The trend of changes in Principal Component Analysis exposed against noise at different levels

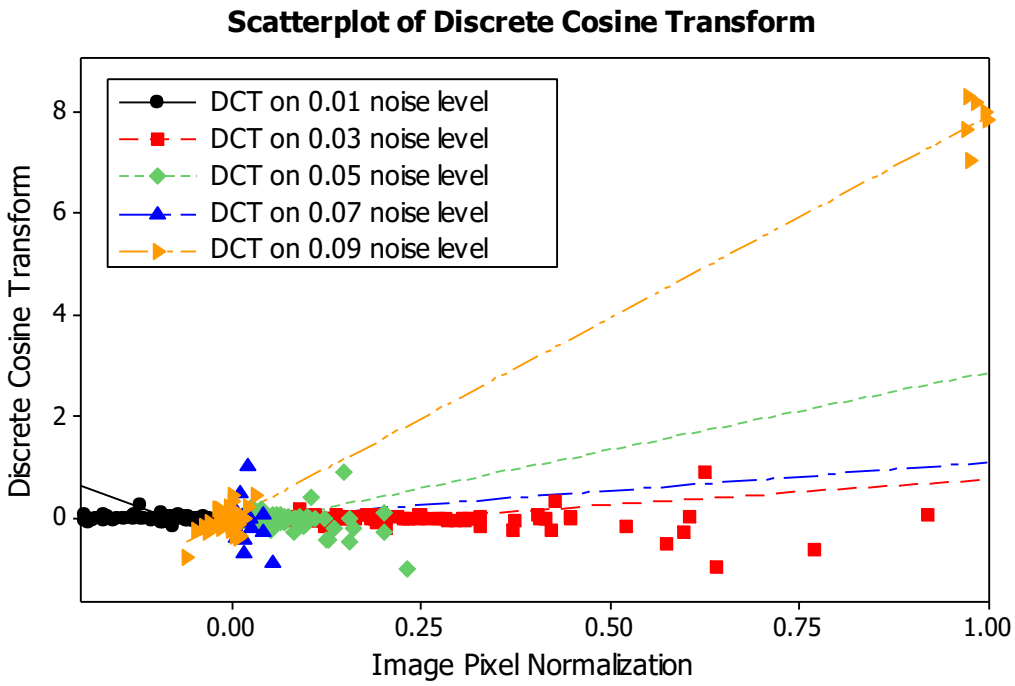


Fig. 11. The trend of changes in Discrete Cosine Transform coefficients exposed to noise at different levels

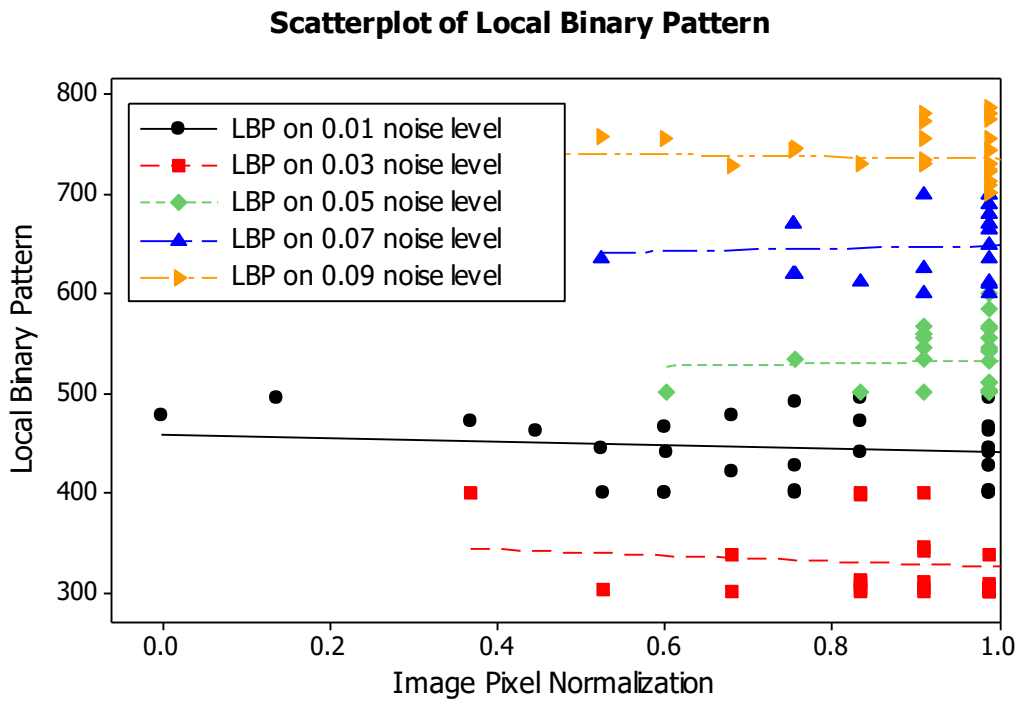


Fig. 12. The trend of changes in Local Binary Pattern coefficients exposed against noise at different levels

After performing a scatter plot and regression analysis, the trend in the correlation between the particular feature extraction quality parameter and noise at different levels was observed. Subsequently, the features with significant correlation between the two elements were chosen. In

this case, they must meet both criteria below to be selected as the model's variable. Those criteria are $R^2 > 80\%$ and p-value < 0.05.

Table 2
t-test Analysis for various Features

Method	Level Noise	*p*	r^2
Principal Component Analysis (PCA)	0.01	0.012	95.5
	0.03	0.033	96.6
	0.05	0.021	90.8
	0.07	0.024	94.2
	0.09	0.013	90.2
Discrete Cosine Transform coefficients (DCT)	0.01	0.022	93.0
	0.03	0.043	99.6
	0.05	0.041	98.7
	0.07	0.014	99.2
	0.09	0.011	98.6
Local Binary Patterns (LBP)	0.01	0.032	90.5
	0.03	0.013	90.6
	0.05	0.029	89.8
	0.07	0.029	84.2
	0.09	0.038	80.2

According to the results in Table 2, the significant correlations to be used in the model are those between all feature parameters and the noise level. It is therefore concluded that all three features are suitable candidates as inputs, given the criteria $p < 0.05$.

3.5 Features Combination

PCA, DCT, and LBP are not redundant feature extractors. PCA primarily represents dominant global variations and reduces dimensionality; DCT represents spatial-frequency information; and LBP describes local facial texture and structural patterns. Their combination provides a broader representation of infant facial pain cues than relying on a single feature descriptor.

It also depicts that several different fusion feature-extraction schemes, including the product of the sum, min, and max rules, are adopted [49, 3, 6]. According to the literature, the sum rule was the most resilient to estimation errors and performed best [49]. Thus, the PCA, DCT and LBP features are fused using the simple sum rule: $\sum PCA + \sum DCT + \sum LBP$. This combination also addresses a limitation of individual descriptors. PCA alone may provide an efficient low-dimensional representation but does not explicitly describe local texture. DCT provides frequency information but does not directly encode local texture patterns. LBP captures local texture effectively but may have difficulty representing larger-scale facial structures when using a small neighbourhood. Feature fusion therefore allows the strengths of the three descriptors to complement one another. The resulting fused features are used as input to the SVM classifier to make the final decision.

4. Support Vector Machine Classifier

SVMs [50] are classifiers designed to discover the best possible hyperplane separator of the two classes. They are extremely efficient in the solution of clustering problems that are non-linearly separable. The main concept consists in projecting (through the use of a mapping function ϕ) the input vectors into a new space of higher dimensions where it is possible to find a hyperplane capable

of creating a linear separation of the two classes. Practically speaking, however, SVMs make use of kernel functions for calculating scalar products in the feature space, without having to specify ϕ. The optimal hyperplane is selected on the basis of its maximum distance from the vectors closest to the border of the separator. The classification of a sample x is given by Equation (18).

$$f(x) = \sum_{i=1}^{N} \alpha_i \, y_i k(x, x_i) + b, \tag{18}$$

where k is the kernel function, xi are the training samples and $yi \in \{-1 + 1\}$ respective class labels, αi and b are the model parameters obtained after training. At test time, f(x) was compared to a threshold θ. All SVM experiments were processed in the same Matrix Laboratory (MATLAB) environment used for the experiments. SVM was implemented using the Ohio State University (OSU) SVM Classifier MATLAB Toolbox developed by Ohio State University. Different types of SVM (Polynomial kernel), SVM (Multi-Layer Perceptron kernel), SVM (Radial Basis Function kernel), and SVM (Linear kernel) are explored in the experiments conducted in this study, based on [9, 3, 6].

5. Experimental results

In this work, a conventional validation [51] scheme is used to assess the reliability of the classification results. The proposed feature vectors are randomly divided into 10 sets, and training is repeated 10 times. Specifically, in the Conventional Validation scheme (one training set and one testing set), 204 segments (144 normal and 60 pain) were used. Out of the features extracted from 204 segments, the features extracted from randomly selected 123 segments (60%) were used for training, and the features extracted from the remaining 81 segments (40%) were used for testing. All the networks were trained and tested using a fusion of PCA, LBP and DCT features. This section evaluates the performance of the proposed method for illumination-invariant detection in terms of recognition accuracy, specificity, sensitivity and error rate under different noise levels. The accuracy, specificity, sensitivity and error are defined by Equations (12), (13), (14) and (15), respectively.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{12}$$

$$Specificit = \frac{TN}{TN+FP} \tag{13}$$

$$Sensitivit = \frac{TP}{TP+FN} \tag{14}$$

$$Error = \frac{FP}{FP+TN} \tag{15}$$

True Positives (TP), True Negatives (TN), False Positives (FP) and False Negatives (FN)

5.1 Accuracy Results

Average accuracy results for the SVM (Polynomial kernel), SVM (MLP kernel), SVM (RBF kernel) and SVM (Linear kernel) classifier using the Conventional Validation scheme with fusion features analysis are depicted in Fig. 13. Average classification accuracies of normal and pathological infant images (pain) are tabulated. Furthermore, the average classification accuracy clearly reveals the consistency of the classifier's results. If accuracy is higher, classification results are consistent, and this also reveals that the learning parameters of the classifiers affect their performance.

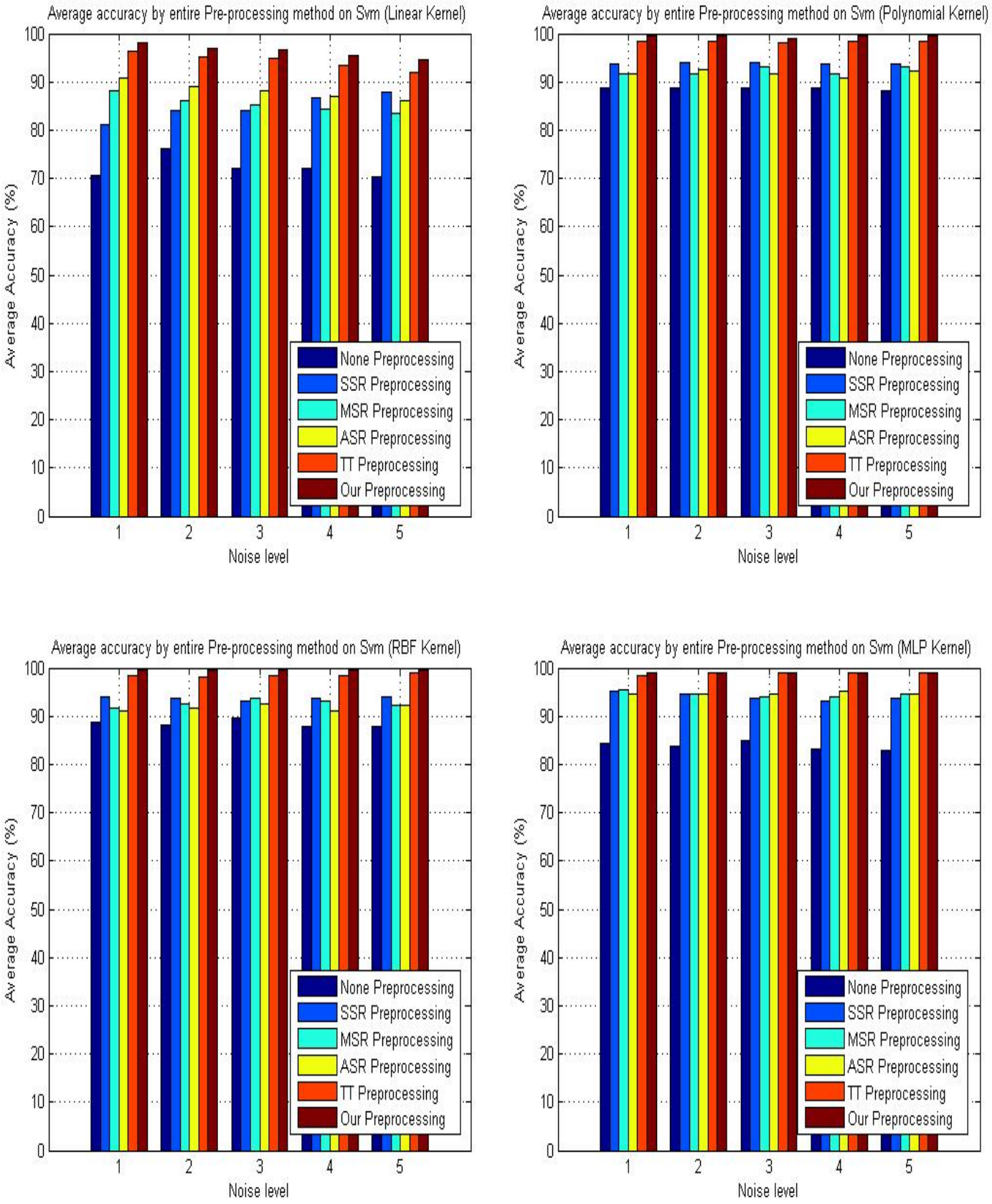


Fig. 13. Average Accuracy by Entire Preprocessing method on Different SVM classifiers

The entire SVM algorithm is trained with different spread factors, and generalization performance is evaluated over 0.5-0.8 and 10-80, respectively. Its effects on the classification performance are analyzed. The maximum classification accuracy was highlighted.

From Fig. 13, the best overall accuracies of 98.01% (OPM at 0.01 noise level), 96.21% (TT at 0.01 noise level), 96.54% (ASR at 0.01 noise levels), 88.22% (MSR at 0.01 noise level), 87.99% (SSR at 0.01 noise level), 76.12% (None at 0.01 noise levels) were obtained using SVM (Linear kernel) classifier. The best overall accuracies of 99.98% (OPM at 0.01 noise level), 98.53% (TT at 0.01 noise level), 94.11% (SSR at 0.01 noise level), 93.13% (MSR at 0.05 noise level), 92.64% (ASR at 0.03 noise level), 88.72% (None) at 0.01 noise levels) were achieved using SVM (Polynomial kernel).

Additionally, the data revealed that the best overall accuracies of 99.70% (OPM at 0.01 noise level), 99.01% (TT at 0.09 noise level), 94.11% (SSR at 0.01 noise level), 93.62% (MSR on 0.05 noise level), 92.64% (ASR at 0.03 noise level), 89.70% (None at 0.01 noise level) were obtained using SVM (MLP kernel). Using SVM (RBF kernel), the best overall accuracies of 99.97% (OPM at 0.01 noise level), 99.01% (TT at 0.01 noise level), 95.58% (MSR at 0.01 noise level), 95.10% (SSR at 0.05 noise level), 95.01% (ASR on 0.03 noise level), 84.80% (None at 0.01 noise level) were attained.

From the above discussion, it has been observed that this proposed method can provide the most discriminating representation of normal and painful infants among other methods. Classification accuracies above 98% are achieved for all SVM algorithms. Finally, the experimental results demonstrate the strength of the proposed method and its potential to detect a pathological condition in an infant from image signals.

5.2 Specificity Results

Specificity Performance results for the SVM (Polynomial kernel), SVM (MLP kernel), SVM (RBF kernel) and SVM (Linear kernel) classifier using the Conventional Validation scheme with fusion features analysis are depicted in Fig. 14. In particular, the specificity of normal and pathological infant images (pain) is tabulated. Specificity refers to a test's ability to correctly identify negative results. As such, higher specificity indicates that the proposed method has a low error rate in the classifier results. The entire SVM algorithm is trained with different spread factors, and generalization performance is evaluated over 0.5-0.8 and 10-80, respectively. Its effects on the classification performance are analyzed. The maximum specificity was highlighted.

From Fig. 14, the best overall specificity of 0.96 (OPM at 0.01 noise level), 0.98 (TT at 0.01 noise level), 0.96 (MSR at 0.01 noise level), 0.91 (ASR at 0.01 noise level), 0.85 (SSR on 0.01 noise levels), 0.85 (None on 0.01 noise levels) was achieved using an SVM (Linear kernel) classifier. Furthermore, the best overall specificity of 0.98 (OPM at 0.01 noise level), 0.96 (TT at 0.01 noise level), 0.95 (MSR at 0.01 noise level), 0.95 (ASR at 0.05 noise level), 0.90 (None at 0.03 noise level), 0.86 (SSR at 0.01 noise level) was obtained using SVM (Polynomial kernel).

Moreover, in the same figure, the best overall specificity of 0.98 (TT at 0.01 noise level), 0.98 OPM at 0.09 noise level), 0.96 (MSR at 0.01 noise level), 0.91 (ASR at 0.02 noise level), 0.85 (SSR at 0.01 noise level), 0.85 (None at 0.01 noise level) was attained using SVM (MLP kernel). Overall specificity of 0.98 (OPM at 0.01 noise level), 0.96 (TT at 0.01 noise level), 0.95 (MSR at 0.01 noise level), 0.95 (ASR at 0.05 noise level), 0.91 (None at 0.03 noise level), 0.87 (SSR at 0.01 noise level) was obtained using SVM (RBF kernel).

Collectively, it can be concluded that this proposed method can provide the most specific representation of normal and painful infants among other methods. Specificity values above 0.85 are achieved across all SVM algorithms. Overall, the experimental results demonstrate the strength of

the proposed method and its potential to detect a pathological condition in an infant from image signals.

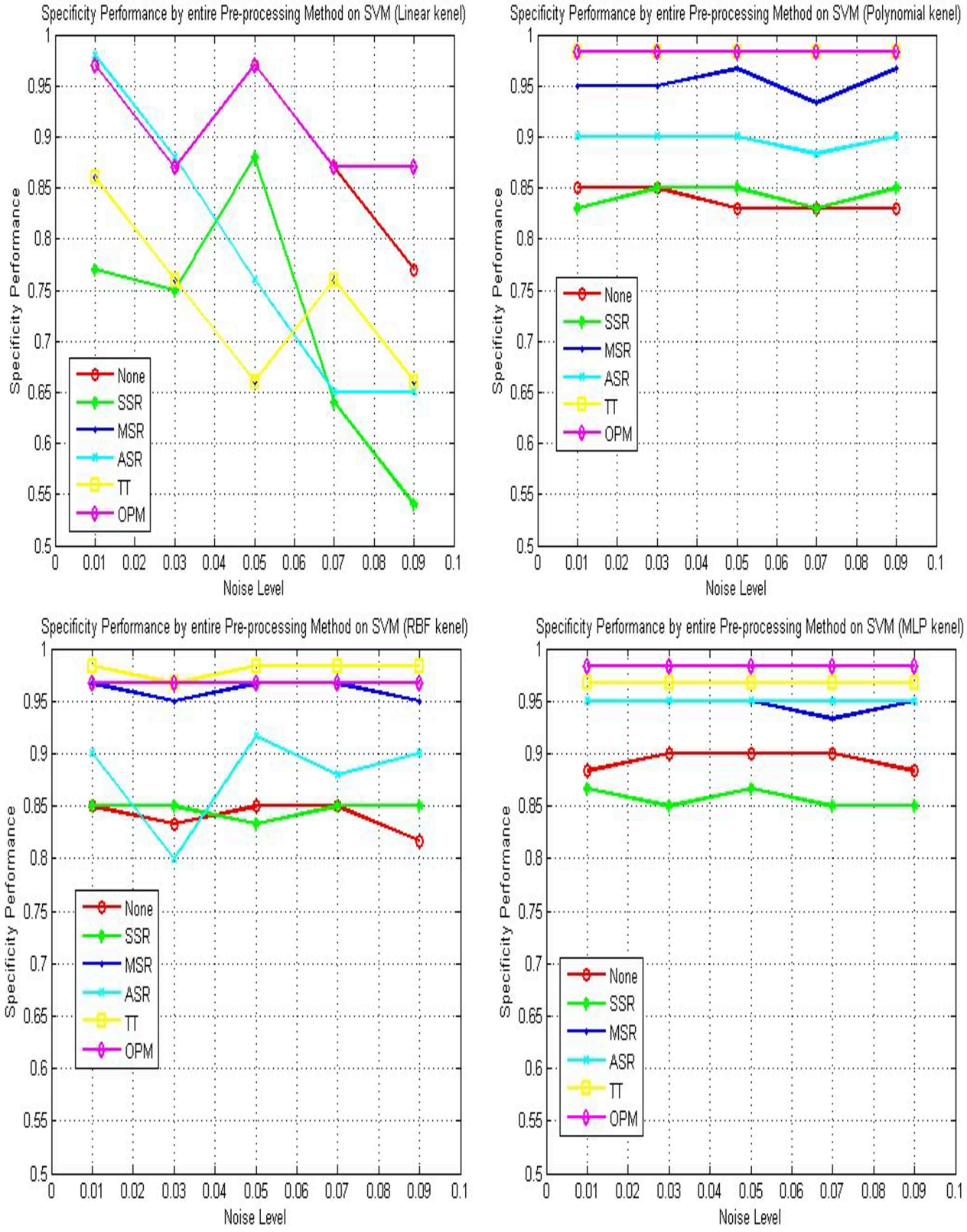


Fig. 14. Specificity of the preprocessing methods using different SVM classifiers.

5.3 Sensitivity Results

Sensitivity results for the SVM (Polynomial kernel), SVM (MLP kernel), SVM (RBF kernel) and SVM (Linear kernel) classifier using the Conventional Validation scheme with fusion features analysis are depicted in Fig. 15. Sensitivity of normal and pathological infant images (pain) is illustrated. The higher sensitivity clearly reveals the lower error of the classifier results. The entire SVM algorithm is trained with different spread factors, and generalization performance is evaluated over 0.5-0.8 and 10-80, respectively. Its effects on the classification performance were analyzed. The maximum sensitivity was highlighted.

From Fig. 15, the best overall sensitivity of 0.88 (OPM at 0.01 noise level), 0.87 (TT at 0.01 noise level), 0.87 (MSR at 0.01 noise level), 0.87 (ASR at 0.01 noise level), 0.80 (SSR at 0.01 noise level), 0.807 (None at 0.01 noise level) was obtained using an SVM (Linear kernel) classifier. Furthermore, the best overall sensitivity of 0.993 (OPM at 0.01 noise level), 0.97 (SSR at 0.01 noise level), 0.96 (TT at 0.01 noise level), 0.937 (ASR at 0.05 noise level), 0.91 (MSR at 0.03 noise level), 0.90 (None at 0.01 noise level) was attained using SVM (Polynomial kernel).

Additionally, the best overall sensitivity of 0.99 (OPM at 0.01 noise level), 0.99 (TT at 0.01 noise level), 0.986 (SSR at 0.01 noise level), 0.958 (MSR at 0.05 noise level), 0.951 (ASR at 0.03 noise level), 0.82 (None at 0.01 noise level) was obtained using SVM (MLP kernel). Finally, the best overall sensitivity of 0.993 (OPM at 0.01 noise level), 0.993 (TT at 0.09 noise level), 0.993 (ASR at 0.01 noise level), 0.97 (SSR at 0.02 noise level), 0.923 (MSR at 0.01 noise level), 0.917 (None at 0.01 noise level) was achieved using SVM (RBF kernel).

It can be concluded that this proposed method can provide the most sensitive representation of normal and painful infants among other methods. Sensitivities above 0.88 are achieved across all SVM algorithms. Overall, the experimental results demonstrate the strength of the proposed method and its potential to detect a pathological condition in an infant from image signals.

5.4 Error Results

Error rates for the SVM (Polynomial kernel), SVM (MLP kernel), SVM (RBF kernel), and SVM (Linear kernel) classifiers using the Conventional Validation scheme with fusion features analysis are tabulated in Table 3. The error rates for normal and pathological infant images (pain) are tabulated. The error rate indicates the proportion of incorrect classifications. Notably, lower error rates yield more precise classification results, indicating that the classifiers' learning parameters affect their performance. The entire SVM algorithm is trained with different spread factors, and generalization performance is evaluated over 0.5-0.8 and 10-80, respectively. Its effects on the classification performance are analyzed. The minimum error was highlighted.

From Table 3, the lowest overall error of 0.02 (OPM at 0.01 noise level), 0.05 (TT at 0.05 noise level), 0.05 (MSR at 0.05 noise level), 0.05 (ASR at 0.05 noise level), 0.05 (SSR at 0.01 noise level), 0.07 (None at 0.07 noise level) was obtained using SVM (Linear kernel).

The lowest overall error of 0.004 (OPM at 0.09 noise level), 0.01 (TT at 0.01 noise level), 0.06 (MSR at 0.09 noise level), 0.07 (ASR at 0.09 noise level), 0.11 (None at 0.03 noise level), 0.05 (SSR at 0.01 noise level) was attained using SVM (Polynomial kernel).

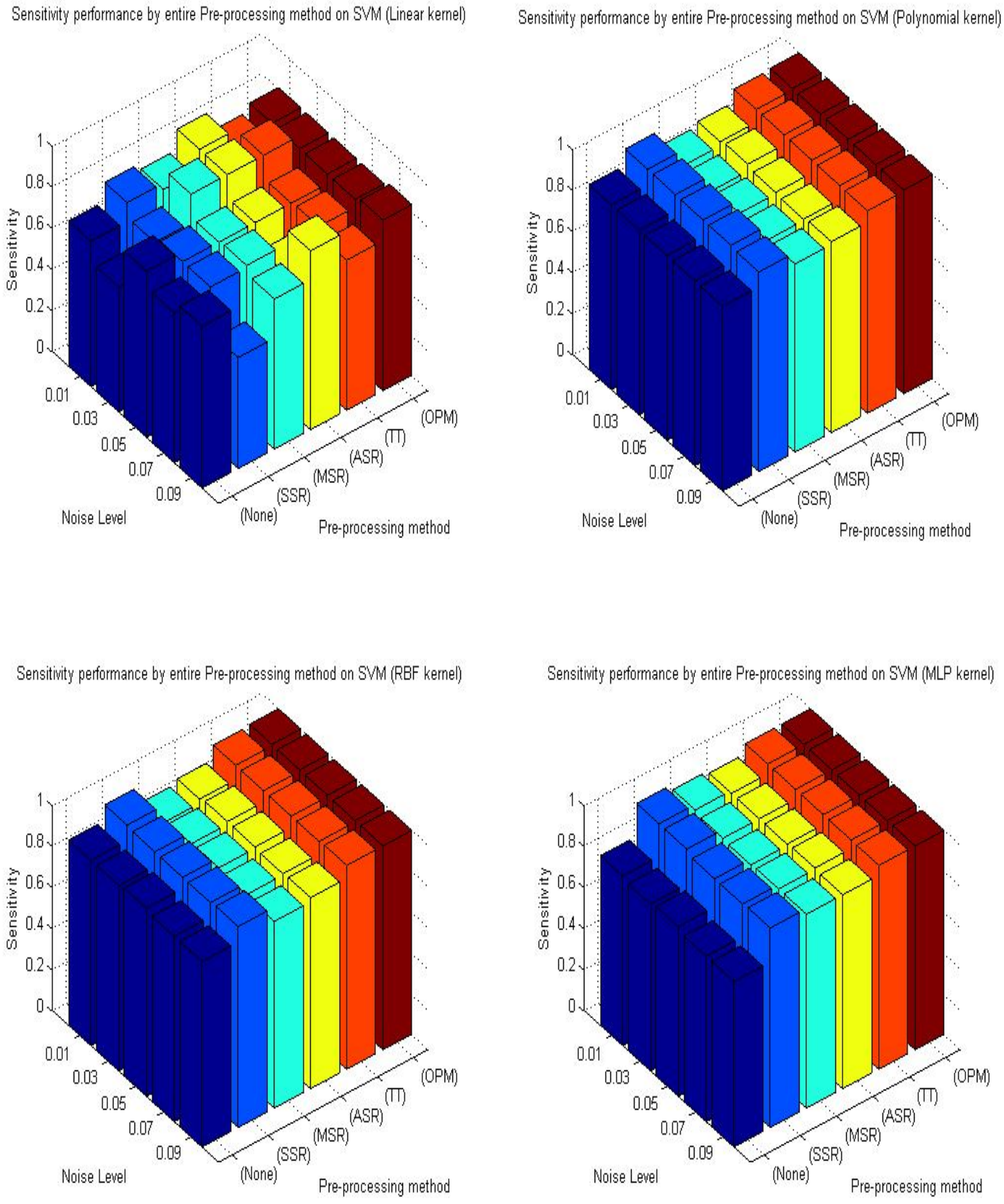


Fig. 15. Sensitivity Performance by Entire Preprocessing method on Different SVM classifiers

Moreover, the lowest error of 0.004 (OPM at 0.01 noise level), 0.01 (TT at 0.07 noise level), 0.06 (MSR at 0.05 noise level), 0.07 (ASR at 0.05 noise level), 0.11 (None at 0.05 noise level), 0.05 (SSR at 0.09 noise level) was obtained using SVM (RBF kernel).
Finally, the lowest overall error of 0.009 (OPM at 0.05 noise level), 0.009 (TT at 0.03 noise level), 0.04 (MSR at 0.01 noise level), 0.04 (ASR at 0.07 noise level), 0.15 (None at 0.01 noise level), 0.04 (SSR at 0.01 noise level) was achieved using SVM (MLP kernel).

Therefore, it can be concluded that this proposed method can provide the lowest error rate representation of normal and painful infants, among other methods. The lowest error rate of 0.004 was achieved with SVM (Polynomial kernel) algorithms. Collectively, the experimental results demonstrate the strength of the proposed method and its potential to detect a pathological condition in an infant from image signals.

Table 3
Error Rates Performance by the Entire Preprocessing Method on Different SVM classifiers

		Classifier Error Rates			
Pre-Processing	Noise Levels	SVM (Linear kernel)	SVM (Polynomial kernel)	SVM (RBF kernel)	SVM (MLP kernel)
None	0.01	0.08	0.11	0.11	0.15
	0.03	0.07	0.11	0.11	0.16
	0.05	0.07	0.11	0.11	0.15
	0.07	0.07	0.11	0.11	0.16
	0.09	0.08	0.11	0.12	0.17
SSR	0.01	0.05	0.06	0.05	0.04
	0.03	0.06	0.05	0.06	0.05
	0.05	0.05	0.05	0.06	0.06
	0.07	0.07	0.06	0.06	0.06
	0.09	0.07	0.06	0.05	0.06
MSR	0.01	0.06	0.08	0.08	0.04
	0.03	0.06	0.08	0.07	0.05
	0.05	0.05	0.06	0.06	0.05
	0.07	0.07	0.08	0.06	0.05
	0.09	0.06	0.06	0.07	0.05
ASR	0.01	0.05	0.08	0.08	0.05
	0.03	0.05	0.07	0.08	0.05
	0.05	0.06	0.08	0.07	0.05
	0.07	0.06	0.09	0.08	0.04
	0.09	0.05	0.07	0.07	0.05
TT	0.01	0.06	0.01	0.01	0.01
	0.03	0.06	0.01	0.01	0.009
	0.05	0.05	0.01	0.01	0.009
	0.07	0.07	0.01	0.01	0.009
	0.09	0.06	0.01	0.009	0.009
OPM	0.01	0.02	0.004	0.004	0.009
	0.03	0.04	0.004	0.004	0.009
	0.05	0.07	0.009	0.004	0.009
	0.07	0.07	0.004	0.004	0.009
	0.09	0.08	0.004	0.004	0.009

6. Conclusion

This study addressed the problem of automated infant pain detection from facial images under varying image-quality conditions, with particular emphasis on illumination variation and noise. The main research gap addressed was the limited attention given to image preprocessing in infant pain recognition. Rather than treating preprocessing as a preliminary image-enhancement step, the study investigated preprocessing as an important component of the overall recognition process. The proposed framework combined histogram remapping, Adaptive Single Scale Retinex (ASR), gamma correction, and contrast equalization to improve image consistency, reduce illumination-related variations, and enhance facial information before feature extraction.

The findings demonstrate that preprocessing has a substantial influence on the quality of the subsequent facial representation. The proposed preprocessing method produced more discriminative representations than the compared preprocessing conditions, particularly when the images were affected by noise. This finding supports the central premise of the study that reliable automated infant pain recognition depends not only on the classifier but also on the quality and stability of the visual information supplied to the classifier. The proposed method achieved maximum classification accuracies of 99.98% using the Polynomial SVM, 99.97% using the RBF SVM, 99.70% using the MLP SVM, and 98.01% using the Linear SVM.

The feature-fusion results further indicate the importance of combining complementary representations. PCA, DCT, and LBP describe different properties of the processed facial images. PCA provides a compact representation of dominant image variations, DCT captures spatial-frequency information through its energy-compaction property, and LBP represents local facial texture and spatial patterns. Their fusion therefore provides a broader description of infant facial pain cues than a single feature representation. The statistical analysis under different noise levels also supported the suitability of all three feature types for inclusion in the fused representation. An important finding is that preprocessing and feature fusion should be considered as interconnected stages rather than independent components. Preprocessing reduces unwanted variation caused by illumination and noise, thereby producing a more consistent facial representation. PCA, DCT, and LBP then exploit this improved representation from complementary perspectives. The sum-rule fusion integrates these representations before classification, allowing the SVM to make its decision from a more comprehensive feature space.

The classification results provide further evidence of the effectiveness of this integrated approach. The proposed method achieved a maximum specificity of 0.98 using the Polynomial SVM. and sensitivity values of up to 0.993 for Polynomial SVM and RBF SVM.

This demonstrates its ability to distinguish between pain and non-pain images while maintaining a high pain-detection rate. The lowest reported classification error was 0.004 using the Polynomial and RBF SVM configurations. This suggests that the fused feature representation contains sufficiently discriminative information to support high-performance nonlinear classification. However, the results should be interpreted within the experimental conditions of this study rather than as evidence of clinical readiness. The experiments were conducted using the Infant COPE database and conventional validation, and the study itself recognizes the need for evaluation using larger samples. Overall, the research demonstrates that robust infant pain recognition requires attention to the complete processing pipeline, from image enhancement to feature representation and classification. The principal contribution is the integration of a noise- and illumination-oriented preprocessing framework with complementary PCA, DCT, and LBP feature fusion for automated infant pain classification. The results show that improving the quality of the input representation and combining complementary feature information can substantially enhance recognition performance.

Nevertheless, further research is required to determine whether the proposed representation remains reliable under larger and more diverse infant populations, uncontrolled lighting conditions, and more realistic clinical environments. Future work should therefore investigate additional preprocessing and feature-extraction techniques, hybrid classifiers, unstructured illumination conditions, and larger datasets to establish a more generalizable and clinically robust automated infant pain detection system.

Acknowledgement
This research was conducted with the support of Universiti Malaysia Perlis (UniMAP) through the Research Acceleration Internal Funding (REACTIF) scheme [Grant Number: UniMAP.RMC/9001-00822(1)].

References

[1] Bonica, J. J. (1979). The need of a taxonomy. Pain, 6(3), 247–248. https://doi.org/10.1016/0304-3959(79)90046-0
[2] Brahnam, S., C. Chuang, F. Y. Shih, and M. R. Slack. "Machine Recognition and Representation of Neonatal Facial Displays of Acute Pain." Artificial Intelligence in Medicine 36, no. 2 (2006): 211–222.
[3] Brahnam, S., C. Chuang, S. S. Randal, and Y. S. Frank. "Machine Assessment of Neonatal Facial Expressions of Acute Pain." Decision Support Systems 43 (2007): 1242–1254.
[4] Brahnam, S., Nanni, L., & Sexton, R. (2007). "Introduction to neonatal facial pain detection using common and advanced face classification techniques". In M. S. Abdel-Wahab & S. Brahnam (Eds.), Advanced computational intelligence paradigms in healthcare - 1 (Vol. 48, pp. 225–253). Springer. https://doi.org/10.1007/978-3-540-47527-9_9.
[5] Brahnam, S., Chuang, C.-F., & Slack, M. R. (2006). "SVM classification of neonatal facial images of pain". In *Computer Analysis of Images and Patterns* (pp. 121–128). Springer.
[6] Brahnam, S., and L. Nanni. "Neonatal Facial Pain Detection Using NNSOA and LSVM." In Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV08), vol. 2, 352–357. Las Vegas.
[7] Nanni, L., S. Brahnam, and A. Lumini. "A Local Approach Based on a Local Binary Patterns Variant Texture Descriptor for Classifying Pain States." Expert Systems with Applications 37, no. 11: 7888–7894.
[8] Abramowitz, M., & Stegun, I. A. (Eds.). (1972). "Inverse circular functions (Section 4.4). In Handbook of mathematical functions with formulas, graphs, and mathematical tables" (9th printing, pp. 79–83). Dover Publications.
[9] Brahnam, S., C. F. Chuang, F. Y. Shih, and M. R. Slack. "SVM Classification of Neonatal Facial Images of Pain." In Fuzzy Logic and Applications, edited by Isabelle Bloch, Alfredo Petrosino, and Andrea G. B. Tettamanzi, 111–115. Lecture Notes in Computer Science 3849.
[10] Pun, K. H., & Moon, Y. S. (2004). "Recent advances in ear biometrics". In Proceedings - Sixth IEEE International Conference on Automatic Face and Gesture Recognition (pp. 164–169). IEEE. https://doi.org/10.1109/AFGR.2004.1301525.
[11] Spence, K., D. Gillies, D. Harrison, L. Johnston, and S. Nagy. "A Reliable Pain Assessment Tool for Clinical Assessment in the Neonatal Intensive Care Unit." Journal of Obstetric, Gynecologic & Neonatal Nursing 34 (2003): 80–86.
[12] Cignacco, E., R. Mueller, J. P. H. Hamers, and P. Gessler. "Pain Assessment in the Neonate Using the Bernese Pain Scale for Newborns." Early Human Development 78 (2004): 115–121.
[13] Lawrence, J., D. Alcock, P. McGrath, S. Kay, S. B. MacMurray, and D. Dulberg. "The Development of a Tool to Assess Neonatal Pain." Neonatal Network 11 (1993): 59–66.
[14] Krechel, S. W., and J. Bildner. "CRIES: A New Neonatal Postoperative Pain Measurement Score. Initial Testing of Validity and Reliability." Paediatric Anaesthesiology 5 (1995): 53–61.
[15] Stevens, B., C. Johnston, P. Petryshen, and A. Taddio. "Premature Infant Pain Profile: Development and Initial Validation." Clinical Journal of Pain 11 (1996): 12–22.
[16] Hummel, P., P. Lawlor-Klean, and M. G. Weiss. "Validity and Reliability of the N-PASS Assessment Tool with Acute Pain." Journal of Perinatology (2009): 1–5.
[17] Hummel, P., 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." Journal of Perinatology 16 (2008): 55–60.

[18] Jobson, D. J., Z. Rahman, and G. A. Woodell. “Properties and Performance of a Center/Surround Retinex.” IEEE Transactions on Image Processing 6, no. 3 (1997): 451–462.
[19] Land, E. H., and J. J. McCann. “Lightness and Retinex Theory.” Journal of the Optical Society of America 61, no. 1 (1971): 1–11.
[20] Leszczynski, M. “Image Preprocessing for Illumination Invariant Face Verification.” Journal of Telecommunications and Information Technology 4: 19–25.
[21] Tan, X., and B. Triggs. “Enhanced Local Texture Sets for Face Recognition under Difficult Lighting Conditions.” IEEE Transactions on Image Processing 19, no. 6 (2010): 1635–1650.
[22] Štruc, V., J. Žibert, and N. Pavešić. “Histogram Remapping as a Preprocessing Step for Robust Face Recognition.” WSEAS Transactions on Information Science and Applications 6, no. 3 (2009): 520–529.
[23] Park, Y. K., S. L. Park, and J. K. Kim. “Retinex Method Based on Adaptive Smoothing for Illumination Invariant Face Recognition.” Signal Processing 88, no. 8 (2008): 1929–1945.
[24] Štruc, V., and N. Pavešić. “Photometric Normalization Techniques for Illumination Invariance.” In Advances in Face Image Analysis: Techniques and Technologies, edited by Y. J. Zhang, 279–300. IGI Global, 2011.
[25] Štruc, V., and N. Pavešić. “Gabor-Based Kernel-Partial-Least-Squares Discrimination Features for Face Recognition.” Informatica (Vilnius) 20, no. 1 (2009): 115–138.
[26] Kittler, J., M. Hatef, R. Duin, and J. Matas. “On Combining Classifiers.” IEEE Transactions on Pattern Analysis and Machine Intelligence 20, no. 3 (1998): 226–239.
[27] Raymond, H. C., H. C. Wa, and M. Nikolova. “Salt-and-Pepper Noise Removal by Median-Type Noise Detectors and Detail-Preserving Regularization.” IEEE Transactions on Image Processing 14, no. 10 (2005): 1479–1485.
[28] Jagadish, H. P. “Robust Fuzzy Median Filter for Impulse Noise Reduction of Gray Scale Images.” World Academy of Science, Engineering and Technology 40 (2010): 635–639.
[29] James, C. C., Y. Chen, and V. Stephen. “A Spatial Median Filter for Noise Removal in Digital Images.” In IEEE Conference Proceedings, 618–623. 2008.
[30] McAndrew, A. An Introduction to Digital Image Processing with MATLAB: Notes for SCM2511 Image Processing 1. School of Computer Science and Mathematics, Victoria University of Technology, 1–264.
[31] Jobson, D. J., Z. Rahman, and G. A. Woodell. “A Multiscale Retinex for Bridging the Gap between Color Images and the Human Observations of Scenes.” IEEE Transactions on Image Processing 6, no. 7 (1997): 965–976.
[32] Hotelling, H. “Analysis of a Complex of Statistical Variables into Principal Components.” Journal of Educational Psychology 24 (1933): 417–441.
[33] Van der Maaten, L. J. P. An Introduction to Dimensionality Reduction Using MATLAB. Technical Report MICC-IKAT 07-07. Maastricht: Maastricht University, 2007.
[34] Turk, M. A., and A. P. Pentland. “Face Recognition Using Eigenfaces.” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 586–591. 1991.
[35] Huber, R., H. Ramoser, K. Mayer, H. Penz, and M. Rubik. “Classification of Coins Using an Eigen Space Approach.” Pattern Recognition Letters 26, no. 1 (2005): 61–75.
[36] Posadas, A. M., F. Vidal, F. de Miguel, G. Alguacil, J. Pena, J. M. Ibanez, and J. Morales. “Spatial-Temporal Analysis of a Seismic Series Using the Principal Components Method.” Journal of Geophysical Research 98, no. B2 (1993): 1923–1932.
[37] Karhunen, J., E. Oja, L. Wang, R. Vigario, and J. Joutsensalo. “A Class of Neural Networks for Independent Component Analysis.” IEEE Transactions on Neural Networks 8, no. 3 (1997): 486–504.
[38] Comon, P. “Independent Component Analysis, a New Concept?” Signal Processing 36 (1994): 287–314.
[39] Olshausen, B. A., and D. J. Field. “Emergence of Simple-Cell Receptive Field Properties by Learning a Sparse Code for Natural Images.” Nature 381 (1996): 607–609.
[40] Rao, K., and P. Yip. Discrete Cosine Transform: Algorithms, Advantages, Applications. Boston: Academic Press, 1990.
[41] Makhoul, J. “A Fast Cosine Transform in One and Two Dimensions.” IEEE Transactions on Acoustics, Speech, and Signal Processing 28, no. 1 (1980): 27–34.
[42] Schwerin, B., and K. K. Paliwal. “Local-DCT Features for Facial Recognition.” In Proceedings of the International Conference on Signal Processing and Communication Systems. Gold Coast, Australia, 2008.
[43] Lee, S. H., Choi, J. Y., Ro, Y. M., & Plataniotis, K. N. (2012). “Local color vector binary patterns from multichannel face images for face recognition”. IEEE Transactions on Image Processing, 21(4), 2347–2353. https://doi.org/10.1109/TIP.2011.2177842.
[44] Ojala, T., Pietikäinen, M., & Harwood, D. (1996). “A comparative study of texture measures with classification based on featured distribution”. Pattern Recognition, 29(1), 51–59. doi.org10.1016/0031-3203(95)00067-4.

[45] He, L., Zou, C., Zhao, L., & Hu, D. (2005). "An enhanced LBP feature based on facial expression recognition". Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 3300–3303. https://doi.org/10.1109/IEMBS.2005.1617182.

[46] Zhang, S., Zhao, X., & Lei, B. (2012). "Facial expression recognition based on local binary patterns and local fisher discriminant analysis". WSEAS Transactions on Signal Processing, 8(1), 21–31. https://doi.org/10.5555/2190016.2190019.

[47] Littlewort, G., Bartlett, M. S., Fasel, I., Susskind, J., & Movellan, J. (2004). "Dynamics of facial expression extracted automatically from video". Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop, 85. https://doi.org/10.1109/CVPR.2004.327.

[48] Lyons, M. J., Budynek, J., & Akamatsu, S. (1999). "Automatic classification of single facial images". IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(12), 1357–1362. doi.org10.1109/34.817413.

[49] Press, William H., Saul A. Teukolsky, William T. Vetterling, and B. P. Flannery. "Section 16.5: Support Vector Machines." In Numerical Recipes: The Art of Scientific Computing. 3rd ed. New York: Cambridge University Press, 2007.

[50] Nakayama, Y. (2022). Support vector machine and optimal parameter selection for high-dimensional imbalanced data. Communications in Statistics – Simulation and Computation, 51(11), 6739–6754. https://doi.org/10.1080/03610918.2020.1813300.

[51] Kohavi, R. "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection." Paper presented at the 14th International Joint Conference on Artificial Intelligence, Montreal, Quebec, Canada, 1995.