Journal of Theoretical and Applied Information Technology
20th September 2014. Vol. 67 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A NOVEL CLASSIFICATION SCHEME FOR NON-INVASIVE HEMOGLOBIN MEASUREMENT METHODS

[1]KUMAR R [2]Dr RANGANATHAN H
[1]Research Scholar, Sathyabama University, Chennai-600119, India
[2]Principal, Rajiv Gandhi College Engineering, Chennai-602117, India
E-mail: [1]biomedkumaran@yahoo.co.in, [2]ranlal@yahoo.com

ABSTRACT

Hemoglobin (Hb) is an imperative constituent of red blood cells (RBC). Assessment of individual physiological status is accomplished by measuring Hb concentration in human blood. The limitations associated with invasive methods are distress during sample collection, wide analysis time and disability to provide real time monitoring. These limitations were addressed by a gamut of non-invasive schemes. The decision scheme for different blood samples associated with non-invasive schemes exhibits intrinsic intricacies. This paper proposes a classification scheme for diverse blood samples by criterion based identification and association to an existing data set. The scheme was implemented and tested on three samples with different testing conditions. Accuracy was identified as the performance metric for this scheme and the results shows that the proposed schemes displays accuracy ranging from 99% to 72% for a range of 0 dB to 10dB Signal to Noise Ratio (SNR).

Keywords: *Hemoglobin (Hb), Red Blood Cell (RBC), Invasive, Non-Invasive, Signal to Noise Ratio (SNR)*

1. INTRODUCTION

The critical requirement of the intensive care environment in hospitals is to monitor and assess the patient's capacity to transport oxygen. The capacity depends on hemoglobin concentration and oxygen saturation. Hemoglobin (Hb) is the protein residing in red blood cells that collects oxygen from the lungs and distributes it to the body's other tissues. Hb is a tetramer, the molecule is composed of four polypeptide chains and each encloses an iron atom known as haem. The chains protect the haem from water; this allows the haem to combine with and release oxygen [2].

Anemia is the dearth of vigorous red blood cells (RBC) flowing in the vascular system. Conventional definition for anemia by WHO is the presence of hemoglobin (Hb) concentration below 12 g/dL for females and below 13 g/dL for males. The studies conclude that anemia affects at least 33% of all patients with cancer, an estimated 65%–95% of all patients with HIV/AIDS, and 70% of all patients with rheumatoid arthritis [1].

Blood Hb is regularly measured primarily to diagnose anemia. The cut-off values are significant to the diagnosis of anemia are broadly defined for dissimilar medical conditions. In critical conditions conservative measurement of hemoglobin entails transference of samples inducing time delayed assessment. The customary method for measuring Hb concentration is by an automated laboratory or point-of-care analyzer, using a venous or capillary blood sample resulting distress and discomposure. The prospective enhancement in patient care with a non-invasive Hb measurement scheme diminishes agony and uneasiness to the patient.

A non-invasive Hb measurement scheme can be implemented by expending the principle of near-infrared spectroscopy in blend with optical image analysis acquired by a charge-coupled device camera positioned at the opposite side of light sources as shown in Figure 1. The scheme initially perceives images of vessels located at the proximal interphalangeal joints and subsequently the absorption pattern of near-infrared radiation is scrutinized to calculate hemoglobin levels. A near-infrared radiation based Noninvasive monitoring scheme to observe Hb levels and oxygen saturation rate can be availed to cater extensive clinical utility.

The optical properties of fluid and skin have been the subject of abundant examinations. It ideally affords noninvasive measurement scope by exploiting optical characteristics. A systematic unswerving model of diffuse reflectance delivers quantitative information with high degree of precision and hypothetically exhibits consistency. This principle was used by diffuse reflectance spectroscopy in the study of hemodynamics.

Journal of Theoretical and Applied Information Technology
20th September 2014. Vol. 67 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

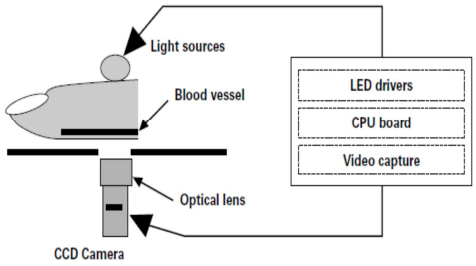

Figure 1: Non-Invasive Hb Measurement Scheme

The objective of this manuscript is to appraise the clinical background and biological background of non-invasive measurement schemes decision making. Section 2 reviews the comprehensive literature on machine learning techniques and ascertains the key design principles for pattern classification techniques. The suitability of these techniques for Hb measurement was studied to uncover an alternative scope. Section 3 proposes an experimental classification scheme and Section 4 evaluates the algorithm's results for different samples with diverse conditions.

2. RELATED WORKS

The process of decision making for future data from previous data by learning is referred as Classification [10]. The previous data applied to the classifier hold the form of feature vectors (collection of attributes). These techniques were founded on three domains namely statistical, machine learning and neural networks [11]. Data capture, data preprocessing and feature extraction are the three steps to establish a classifier [10]. The corroding factors on data are insufficient collection, data miss and noisy data [16]. The probability of a specific case affiliated to precise class can be calculated from prior information of the case. Knowledge of all prior probabilities for each class will enable to realize a perfect classifier. The prior knowledge of all possible classes makes this hypothesis as ideal and unrealizable [12].

Restructuring the natural information is derived from various parameter estimation approaches to discern class conditional pdf is referred as parametric approach. Estimation of pdf by employing meagre number of neighboring samples leads to non-parametric approaches [13]. Parametric methods identify a mathematical model to accommodate the data and performs complete adherence to the chosen model. The possession of single local maximum or peak inhibits these methods to manage multi modal densities. Non parametric methods assess the applied data without prior knowledge of data.

True class labelling of cases and utilization of the same for classifier training is referred as supervised learning. Unsupervised learning also referred as clustering methods assigns classes on the basis of natural groupings [14]. The primary requirement of the classification process is to exhibit well-balanced classification ranging from intricate patterns to simple patterns [15]. The performance metric for classification process is referred as classification error rate and defined as the percentage erroneous pattern assignment [2]. Another bench mark for a classifier is comprehensibility defined as ability to deliver inferable and reliable [2]. The selection criterion for a classifier can be in line with the proposition from No Free Lunch Theorem [2].

Maximum Likelihood Estimation is a statistical and parametric data fitting model and exhibits mandatory data modelling applicable for nonlinear data [2]. A nonlinear modelling possessing the capacity to measure the ability of the model to fit data with low processing time [2]. Linear discriminant functions (LDAs) split the solution space in to smaller domains by using lines, planes and hyper planes. The classification is enacted by localizing the hyper plane position corresponding to the solution [2]. LDAs cannot handle multi-dimensional data [2]. The demerits of the LDAs are addressed by Artificial Neural Networks (ANN) [2].

Support Vector Machines (SVMs) employs the yield of LDAs to an extensive array of classification problems. SVMs map the nonlinear separable data to enlarged dimension and enabling it for hyper plane separation. A kernel function is used to accomplish preprocessing operation [2]. The ability to fit the data aptly is the main advantage of this technique and largely utilized for pattern recognition problems [2]. Prior knowledge of system causality empowers an approach referred as Bayesian nets to conglomerate conditional probabilities for extracting new information. The net components are mutually influencing each other [2].

A Noninvasive Hb measurement system comprises of a noninvasive Hb sensor and Hb meter was proposed in [5]. Intraclass Correlation Coefficients (ICCs) was suggested as performance measure to assess the correlation between the proposed system and an automated hematology analyzer. A study was performed to compare the accuracy of four bedside methods for Hb assessment namely non-invasive and continuous Hb

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

measurement with Pulse Co-Oximetry (SpHb), arterial blood measurement by Satellite Co-Oximetry (HbSat), HemoCue arterial (HcueArt) and laboratory hematology analyzer (LHA) and concluded HcueArt is closest to LHA even when Hb concentrations change rapidly [6]. A quantitative assessment scheme based on light scattering properties to measure Hb and melanin content of skin [7]. The scheme was developed based on diffuse reflectance spectra and absorption coefficient of skin was proposed as performance measure.

An Hb status measurement mechanism to observe advancement of dengue infection based on linear autoregressive moving average with exogenous input (ARMAX) was proposed in [8]. The scheme appraised three models ARMAX model order selection criteria namely Final Prediction Error (FPE), Akaike's Information Criteria (AIC) and Lipschitz number. It demonstrates that Lipschitz number has better accuracy compared to FPE and AIC.

The performance of three noninvasive Hb assessment schemes namely Astrim, Sysmex and Kobe were evaluated in [9]. These schemes measures Hb values and White Blood Cell (WBC) counts. The coefficient of correlation for healthy individuals and hematologic patients was measured and identified Ast-Hb as the best scheme for continuous Hb monitoring [9].

The absorption and scattering coefficients measurement scheme for scrambled optical media was implemented in [10]. It provides Hb concentration and tissue Hb oxygen-saturation (StO_2). These values hold superior accuracy over blood gas analyzer measurements. A PhotoPlethysmoGraphy (PPG) based non-invasive Hb detection using Field Programmable Gate Array (FPGA) based embedded system was implemented in [4]. This non-invasive system was designed to measure blood volume changes in the micro-vascular bed of tissue. The set up can able to monitor PPG waves for three different wavelengths and revealed a high degree of endurance against subject movements. A 3-axis accelerometer was incorporated in the system to measure movement artifacts.

Using reflectance spectroscopy and stochastic photon propagation model a method to determine Hb was proposed in [3]. This model was used to study palpebral conjunctiva and owns the distinction of first biological significant model. The technique delivers Hb levels and blood oxygenation variations. A patterned structure to quantify Hb concentration, oxygen saturation and pulse in perioperative environment was presented in [2]. The methodology was implemented to provide online patient monitoring and swift analysis with minimum risk of infection.

A summary of the works and contemporary trends of research in noninvasive total hemoglobin measurement was studies in [2]. The evaluation of prospective techniques was performed on optic-acoustic spectroscopy, spectrophotometric imaging, diffuse reflectance spectroscopy, transcutaneous illumination, electrical admittance PlethysmoGraphy (PG), and PPG. The techniques were subjected to performance metrics that includes technological performance, relative benefits of each approach, potential instrumentation design considerations.

The related works study discloses a requisite to develop a classification scheme to identify and associate a sample to a sample space of similar characteristics. It also highlights the need to develop a robust algorithm to discern datasets with little correlation. This paper proposes a classification algorithm to measure Hb level accuracy for diverse samples. It implements a comprehensible assessment scheme to function in various data corruptive environments.

3. PROPOSED SCHEME

The samples of scattering and absorption coefficient from 760nm and 850nm wavelength are acquired. The database already consists of samples of these two parameters collected during the training phase for different range of values of Hb. The training phase samples were collected by varying the inspiration oxygen per cent (Fio_2) level in values of 5%, 10%, 15%, 305 and 90% on the subjects. These samples were trained using a classifier and the Hb range was classified. The categorization scheme is presented in Figure 2. The Scattering and Absorption coefficients were given in Table 1, Table 2, Table 3 and Table 4.

Table 1: Absorption Coefficients For Subject 1

ABSORPTION COEFFICIENTS – SUBJECT 1
0.16, 0.17, 0.18, 0.2, 0.2, 0.2, 0.21, 0.22
0.23, 0.23, 0.24, 0.24, 0.25, 0.25, 0.25, 0.25, 0.25
0.26,0.26,0.27,0.27,0.27,0.27,0.273,0.276,0.278, 0.28,0.29
0.17, 0.268, 0.168, 0.166, 0.166, 0.164, 0.164, 0.162, 0.15, 0.15
0.16, 0.17, 0.17, 0.17, 0.16, 0.16, 0.16, 0.16, 0.15, 0.156

Journal of Theoretical and Applied Information Technology
20th September 2014. Vol. 67 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2: Absorption Coefficients For Subject 2

ABSORPTION COEFFICIENTS – SUBJECT 2
0.14, 0.14, 0.14, 0.142, 0.144, 0.146, 0.148, 0.15, 0.157
0.15, 0.151, 0.151, 0.152, 0.153, 0.153, 0.154, 0.154, 0.155, 0.155
0.15, 0.15, 0.15, 0.152, 0.154, 0.156, 0.158, 0.158, 0.158, 0.16
0.15, 0.15, 0.15, 0.15, 0.15, 0.15, 0.15, 0.15, 0.15, 0.16
0.154, 0.152, 0.15, 0.148, 0.148, 0.147

Table 3: Scattering Coefficients For Subject 1

7, 7, 7, 7.1, 7.2, 7.3, 7.3, 7.4, 7.4,
7.5, 7.5, 7.6, 7.6, 7.7, 7.9, 8, 8.1, 8.2
7.8, 8, 8, 8, 8, 7.9, 8, 7.9, 8, 7.9, 8, 7.8, 7.8
7.8, 7.6, 7.5, 7, 7, 7,7, 7,7.1, 7.2
7.1, 7.2, 7.3, 7.3, 7.3, 7.3, 7.3, 7.3, 7.3, 7.3

Table 4: Scattering Coefficients For Subject 2

6,6,6,6,6,6.1,6.1,6.2,6.2,6.2,6.2,
6.2, 6.2, 6.2, 6.2, 6.2, 6.2, 6.2, 6.2, 6.2, 6.2
6.2, 6.2, 6.1, 6.1, 6.1, 6.1, 6.1, 6, 6, 5.9
6.2, 6.2, 6.1, 6.1, 6, 5.9, 5.8, 5.2, 5.7, 5.7
5.7, 5.7, 5.6, 5.6, 5.6, 5.5, 5.4, 5.3, 5.2, 5.2, 5.2

3.1 Pseudo Code

The pseudo code for reading coefficients is given below and Table 5 describes the operation carried out by the pseudo code.

```
j=1;
for i=1 to 4
    read a[i];
    i=i+1;
```

Table 5: Pseudo Code Operation

i=1	Refers to the 1st value of scattering coefficient from 760nm wavelength laser is read
i=2	Refers to the 1st value of scattering coefficient measured from 860nm.
i=3	Refers to the 1st value of absorption coefficient measured from 758nm
i=4	Refer to the 1st value of absorption coefficient measured from 850 nm

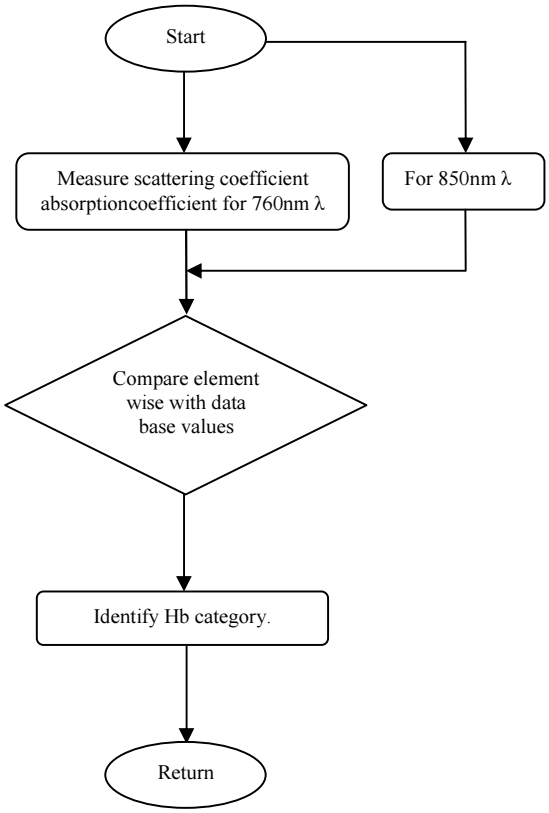

Figure 2: Hb Categorization Scheme

3.2 Hb CLASSIFIER ALGORITHM

The proposed scheme uses the following denotions. The objective of the scheme is to assign class label to unknown samples 'U'.

1. 'U' denotes Unknown samples to be assigned different classes.
2. 'T' denotes Training set containing the training samples (scattering and absorption)

$$T_1 = \{t_{1,1}, t_{1,2}, \dots \dots t_{1,n}\}$$
$$T_2 = \{t_{2,1}, t_{2,2}, \dots \dots t_{2,n}\}$$
$$T_m = \{t_{m,1}, t_{m,2}, \dots \dots t_{m,n}\}$$

3. Feature $t_{i,n}$ be the class label of T_i.
4. 'm' denotes number of training samples.
5. 'n' denotes number of features describing each sample.
6. 'k' denotes number of nearest neighbors to the determined.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Step 1:
Array a[m][2]

i. 'M' represents the rows containing scattering and absorption data regarding 'm' training samples. The first column represents the Euclidean distance between 'U' and that row's training samples
ii. The second column refers to the Sample. The second column refers to that training samples index.
iii. The index needs to be saved, since when sorting the array (according to Euclidean distance), there need to be some method to determine to which training set the Euclidean distance refers.

Step 2:
for i = 1 to m
Step 3:
a[i] [1] = Euclidean_distance(U, T_i)
Step 4:
a[i] [2] = i
In this step the index is saved, as rows will be sorted later.
Step 5:

i. Sort the row of 'a' by their Euclidean distances saved in In this step sorting is done in ascending order

Step 6:
Array b[k][2]

i. The first column holds the distinct class labels of the k-nearest neighbors. The second column holds their respective counts. In the worst case, each k-nearest neighbor will have a different class label, hence there is a need to allocated space for k class label

Step 7:
for i = 1 to K do
Step 8:
If class label $t_{a[i][2]}$ already exists in array 'b' then perform step 9.
Step 9:
Find that class labels row in array 'b' and increment its count
Step 10:
Else add the class label into the next available row of array 'b' and increment its count;}
Step 11:

i. Sort array 'b' in descending order
ii. This sorting is done from class label with largest count down to that with smallest count.

Step 12:
Return the most frequent class label of the k-nearest neighbors of 'U' is returned as the class prediction

The Euclidean distance is defined as

$$d(i,j) = \sqrt{(x_{i1}-x_{j1})^2 + (x_{i2}-x_{j2})^2 + \ldots\ldots + (x_{in}-x_{jn})^2} \quad (1)$$

Where
$i = (x_{i1}, x_{i2}, \ldots x_{in})$ and $j = (x_{j1}, x_{j2}, \ldots x_{jn})$ are two n-dimensional data objects.
Alternately, Manhattan distance defined as

$$d(i,j) = |x_{i1}-x_{j1}| + \cdots + |x_{in}-x_{jn}| \quad (2)$$

Minkowski distance

$$d(i,j) = \left(|x_{i1}-x_{j1}|^q + |x_{i2}-x_{j2}|^q + \cdots + |x_{in}-x_{j1}|^p\right)^{1/p} \quad (3)$$

can be used.

4. RESULTS

The performance of the proposed scheme was assessed by applying for samples Hb1, Hb2 and Hb3. Each sample under different SNRs encompassing four values 0dB, 1dB, 5dB and 10dB. This paper identified accuracy as a performance measure for the methodology. Table 1 presents classified results for Hb1, Hb2 and Hb3 for different SNRs. The Plots SNR (db) versus Accuracy (in %) for Hb1, Hb2 and Hb3 were shown in Figure 3, Figure 4 and Figure 5 respectively.

Table 6: Classified Result for Hb1, Hb2 and Hb3 At Varying SNR Values

Category	SNR			
	0 dB	1 dB	5 dB	10 dB
Hb1	(396, 4, 0, 99%)	(395, 3, 2, 98.75%)	(389, 9, 2, 97.25%)	(395, 1, 4, 98.75%)
Hb2	(95, 304, 1, 76%)	(88, 307, 5, 76.75%)	(96, 303, 1, 75.75%)	(90, 307, 3, 76.75%)
Hb3	(91, 8, 301, 75.25%)	(90, 3, 307, 76.75)	(62, 2, 336, 84%)	(29, 4, 367, 91.75%)

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

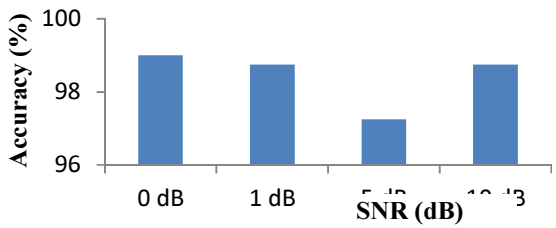

Figure 3: SNR Vs Accuracy Plot for Hb1

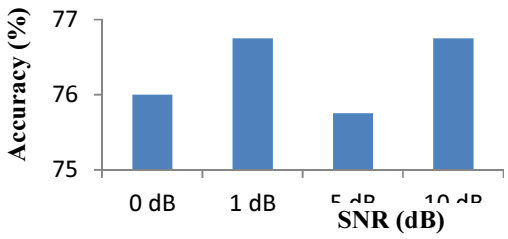

Figure 4: SNR Vs Accuracy Plot for Hb2

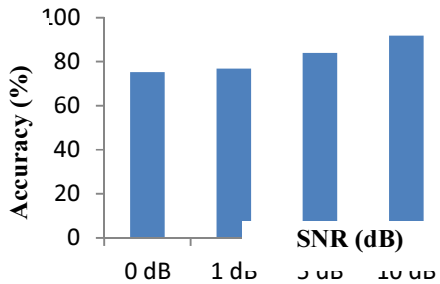

Figure 5: SNR Vs Accuracy Plot for Hb3

If acceptable accuracy is specified as a variable, then the (ROC) Region of convergence curve for the two classifier schemes can be determined. It is inferred that for larger region of operation (ROO), the Hg1 support is good even at lower SNR.

5. CONCLUSION

This paper tenders a method to classify and labels data samples in line with a training set. This implementation facilitates physicians to match clinical profile of a new patient to those of his or her previous patients. The emphasis of proposed scheme on local information consents it to display adaptability. This characteristic can be exploited to use in an online learning system. This methodology can be augmented to the predictor's delinquent realm, its clinical locale and development history to identify unusual versions of the Hb molecule referred as Hb variants ensuing from genetic mutations.

REFERENCES:

[1] John W. McMurdy, Gregory D. Jay, Selim Suner, Gregory Crawford, "Noninvasive optical, Electrical and Acoustic Methods of Total Hemoglobin Determination", *Reviews on Clinical Chemistry*, 2008, pp. 264-272.

[2] Juliette Cunningham, "A k-Nearest Neighbor Prediction Algorithm for Unusual Haemoglobin Variants", *MSc Thesis*, 2007.

[3] U. Timm, E. Lewis, D. McGrath, J. Kraitl and H. Ewald, "LED Based Sensor System for Non-Invasive Measurement of the Hemoglobin Concentration in Human Blood", *Proceedings of International Conferences in Biomedical Engineering*, 2008, pp. 825–828.

[4] Chetan Sharma, Sachin Kumar, Anshul Bhargawa, Shubhajit Row Chowdhury, "Field Programmable Gate Array Based Embedded System for non-invasive estimation of Hemoglobin in blood using Photoplethysmography", *International Journal on Smart sensing and Intelligent Systems*, 2013, Vol.6, No. 3, pp. 1267-1282.

[5] Moon Jung Kim, Quehn Park, Myung Hee Kim, " Comparison of the accuracy of Noninvasive Hemoglobin Sensor (NBM-200) and Portable Hemoglobinometer (Hemocue) with an Automated Hematology Analyzer (LHA500) in Blood Donor Screening", *Annals of Laboratory Medicine*, 2013, pp. 261-268.

[6] B. Giraud, D. Frasca, B. Debaene, O. Mimoz, "Comparison of haemoglobin measurement methods in the operating theatre", *British Journal of Anesthesia,* July 2013, pp. 1-9.

[7] George Zonios, Julie Bykowski, Nikiforos Kollias, "Skin Melanin, Hemoglobin, and Light Scattering Properties can be Quantitatively Assessed In Vivo Using Diffuse Reflectance Spectroscopy", *The Society for Investigative Dermatology,* 2001, pp. 1452-1457.

[8] H. Abdul Rahim, F. Ibrahim, M. N. Taib, R. Abdul Rahim, Y. Mad Sam, "Monitoring

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Haemoglobin Status in Dengue Patients using ARMAX Model", *IEEE,* 2008.

[9] D. M. Hueber, M. A. Franceschini, H. Y. Ma, Q. Zhang, J. R. Ballesteros, S. Fantini, D. Wallace, V. Ntziachristos, B. Chance, "Non-invasive and quantitative near-infrared haemoglobin spectrometry in the piglet brain during hypoxic stress, using a frequency-domain multidistance instrument", *Institute Of Physics Publishing Physics in Medicine and Biology,* 2001, pp. 41-62.

[10] Duda, RO, PE Hart, DG Stork, "Pattern Classification" *Wiley-Interscience*, 2000.

[11] King, R. D., C. Feng, A. Sutherland, "Statlog - Comparison of Classification Algorithms on Large Real-World Problems." *Applied Artificial Intelligence*, 1995, pp. 289-333.

[12] Chan-Yun Yang, Jui-Jen Chou, "A Comparative Evaluation Approach for the Classification of Rotifers with Modified Non-Parametric KNN" *Image and Vision Computing*, 2005,

[13] R. J. Henery, "Classification", Machine *Learning, Neural and Statistical Classification*, 1994, pp. 6.

[14] R. J. Henery, "Methods for Comparison," *Ellis Horwood Series in Artificial Intelligence,* 1995, pp. 107.

[15] I. J. Myung, "Tutorial on Maximum Likelihood Estimation," *Journal of Mathematical Psychology*, 2003, pp. 90.

[16] Oleg Kim, John McMurdy, Gregory Jay, Collin Lines, Gregory Crawford, Mark Alber, "Combined reflectance spectroscopy and stochastic modeling approach for noninvasive hemoglobin determination via palpebral conjunctiva", *Physiological Reports,* 2014, pp. 1-14.