INTERNATIONAL JOURNAL OF HEALTH PROFESSIONS
Volume 2, Issue 1, 2015, Pages 4–15, ISSN 2296-990X, DOI: 10.1515/ijhp-2015-0005

DE GRUYTER OPEN

Comparison of Supervised-learning Models for Infant Cry Classification

Vergleich von Klassifikationsmodellen zur Säuglingsschreianalyse

Tanja Fuhr[1], Henning Reetz[2], Carla Wegener[1]

[1]Hochschule Fresenius, Fachbereich Gesundheit & Soziales, Limburger Str. 2, 65510 Idstein, GERMANY, tanja.fuhr@hs-fresenius.de

[2]Goethe Universität Frankfurt am Main, Institut für Phonetik, Senckenberganlage 31, 60325 Frankfurt, GERMANY

Received 14 November 2014, accepted 22 January 2015, available online 20 April 2015

Abstract

Cries of infants can be seen as an indicator for several developmental diseases. Different types of classification algorithms have been used in the past to classify infant cries of healthy infants and those with developmental diseases. To determine the ability of classification models to discriminate between healthy infant cries and various cries of infants suffering from several diseases, a literature search for infant cry classification models was performed; 9 classification models were identified that have been used for infant cry classification in the past. These classification models, as well as 3 new approaches were applied to a reference dataset containing cries of healthy infants and cries of infants suffering from laryngomalacia, cleft lip and palate, hearing impairment, asphyxia and brain damage. Classification models were evaluated according to a rating schema, considering the aspects accuracy, degree of overfitting and conformability. Results indicate that many models have issues with accuracy and conformability. However, some of the models, like C5.0 decision trees and J48 classification trees provide promising results in infant cry classification for diagnostic purpose.

Abstract

Verschiedene Klassifikationsverfahren konnten bereits zeigen, dass es möglich ist, zwischen gesunden Säuglingsschreien und pathologischen Säuglingsschreien zu unterscheiden. Bislang fehlte jedoch ein systematischer Vergleich der verschiedenen Ansätze. Für diesen Artikel wurden in einer systematischen Literatursuche 9 Klassifikationsmodelle identifiziert, die bereits in der Säuglingsschreiforschung genutzt wurden. Zusammen mit drei weiteren, bislang ungenutzten Ansätzen, wurden die Schreie von gesunden Säuglingen sowie von Säuglingen mit Laryngomalazie, Lippen-Kiefer-Gaumenspalte, Hörstörung, Sauerstoffmangel und Hirnschädigung anhand ihrer akustischen Parameter klassifiziert. Die Leistungsfähigkeit aller Modelle wurde mittels eines standardisierten Schemas nach Genauigkeit, Überanpassung an den Trainingsdatensatz und Nachvollziehbarkeit des Verfahrens bewertet und verglichen. Die Ergebnisse zeigen, dass einige der Modelle Schwächen in der Genauigkeit und Nachvollziehbarkeit aufweisen. Jedoch erzielen Modelle wie die C5.0 und J48 Entscheidungsbäume vielversprechende Ergebnisse, die das Erkennen des jeweiligen Störungsbildes am Schrei mit einer hohen Genauigkeit ermöglichen.

Keywords
supervised-learning models – infant cry – developmental disorders – classification

Keywords
Klassifkationsmodelle – Säuglingsschrei – Entwicklungsstörung

1. INTRODUCTION

The infant cry is one of the earliest capabilities for human communication. In their early stage of life, infants communicate pain, discomfort or desires by crying. Infant cry research in general is an interdisciplinary field of research covering, amongst others, medical and biological aspects like anatomy, physiology and pathology, physical and mathematical aspects like acoustics and signal analysis as well as linguistic aspects like phonetics. Findings of infant cry research are relevant for various health professions like nurses, midwifes or speech therapists and medical professions like pediatricians by helping to interpret the infant cry in order to recognise an infant's needs or health state.

For about six decades, researchers of various scientific professions have been exploring acoustic features of the infant cry, like fundamental frequency, intensity or resonance frequencies. Many studies indicate that the infant cry is suited to identify an infant's mood, for example, hunger, tiredness or discomfort. Apart from that, research has shown that the infant cry may be an early indicator for the infant's health state. Differences in acoustic parameters have been related to central nervous system insults like brain haemorrhage or asphyxia (Lester et al., 2002; Corwin et al., 1992; Blinick et al., 1971; Nugent et al., 1996; Michelsson et al., 1977; Verduzco-Mendoza et al., 2012), to disturbances of the vocal neuromuscular maturation (Lind et al., 1967; Golub and Corwin, 1982; Fort and Manfredi, 1998), to various developmental disorders like hearing impairment (Möller and Schönweiler, 1999; Arch-Tirado et al., 2004; Varallyay, 2007; Etz et al., 2012) or autism (Esposito et al., 2013; Sheinkopf et al., 2012), to brain disorders (Sirvio and Michelsson, 1976; Fisichelli and Karelitz, 1966; Wasz-Hockert et al., 1968; Karelitz and Fisichelli, 1962) and to genetic defects like the Down syndrome (Fisichelli and Karelitz, 1966; Lind et al., 1970), Krabbe's disease (Thoden and Michelsson, 1979) and the cri-du-chat syndrome (Wasz-Hockert et al., 1968; Vuorenkoski et al., 1966).

Inspired by these promising achievements and boosted by the growing capabilities of information technology, much research has been conducted to find approaches for automatically predicting an infant's state of health based on acoustic infant cry features. Several classification model approaches originating from mathematical or computer science disciplines were applied to the infant cry. Classification model algorithms — also called supervised-learning algorithms — are trained on a training dataset containing predictor variables (e.g., acoustic features of an infant cry) and for which the classification result is known (e.g., the state of health of an infant). After the training phase has finished, the classification model can be used to classify data for which the result is not known. By this, an infant's state of health could be predicted by applying a classification model on the acoustic features measured for the infant's cry.

Results reported for training such classification models are promising. However, there is no consensus on which classification models are best suited for infant cry classification. Available reports in the literature are not comparable as models were trained on different data. This article aims to provide a systematic comparison of various models used for infant cry classification. As known to the authors, this is the first systematic comparison of infant cry classification models.

The remaining article is structured as follows. Section 2 introduces the structure of the systematic literature search and the classification model review. Section 3 presents the major results of this review. Models and model ratings are provided. Finally, Section 4 interprets the ratings and provides recommendations for future use of classification models in infant cry research.

2. METHOD

A *systematic classification model comparison* was conducted with a literature search aiming to identify all possible classification models and to compare them systematically. For this purpose, the model review consisted of two major parts, a literature search and the systematic model rating. The literature itself was not reviewed (as would be done in systematic literature reviews), but the *classification models* used were identified. Next, the classification models were applied to a common reference dataset and rated systematically.

The model review was divided into two phases. First, a systematic literature search was conducted to identify which classification models have already been used in infant cry research (Section 2.1). Second, models were compared and rated on their performance in infant cry classification in a model review (Section 2.2). Sections 2.3 and 2.4 describe the dataset used as reference for the model review.

2.1. Systematic Literature Search

For the literature search, the research question as well as inclusion and exclusion criteria were specified.

For the literature search and the model review, the following major research question was stated: 'Which supervised-learning classification model is best suited to classify infant cries according to their state of health'? This main research question was split into two subquestions: (a) What supervised-learning classification models have already been used in infant cry research? and (b) Which one of these models is best suited to classify infant cries according to their state of health?

Inclusion and exclusion criteria were defined. Both sets cover criteria for the publication as well as for the model approach itself. For publications to be included, they had to describe the application of classification models in infant cry research and they had to be written in English or German. The described modelling approach had to be applicable to metric explanatory variables and to nominal predictor variables and had to be robust against non-normally distributed data.

Publications were excluded if they did not provide a sufficient description of the modelling approach in order to understand how the approach worked.

For the literature search, seven online libraries and indexes covering the research field were selected: ACM

digital library (ACM, 2014), DBLP (DBLP, 2014), IEEE XPlore (IEEE, 2014), SSG (GBV, 2014), DIMDI (DIMDI, 2014), Medpilot (Deutsche Zentralbibliothek für Medizin, 2014) and Web of Knowledge (Thomson Reuters, 2014). All data sources were searched for articles describing the application of classification models in infant cry research.
For the search strategy, the search term was composed to cover major keywords relevant for the research question, as well as synonyms and related terms as well as broader terms and narrower terms in order to find all relevant literature. Based on these terms, the following search string was composed:

(classif* OR predict* OR forecast*
OR "machine*learning" OR "supervised*learning")
AND (model* OR algorithm* OR approach*)
AND (infant* OR baby OR babies OR newborn* OR
neonate*)
AND (cry* OR cries)

Applying this search string to the libraries sometimes required adaption to the specific search engine (e.g., some libraries used the symbols '&&' instead of the keyword 'AND'; others did not support wildcards like '*', which required to write down all possible morphological variants of a word, e.g., 'predictive', 'prediction', 'predictions', 'predicting' instead of 'predict*').
After defining the setting, the literature search was conducted. The result sets of articles from each data source were collected in one central bibliography. The articles that were gathered by the initial search were scanned (title, abstract) and obviously irrelevant articles were excluded. In addition, duplicates were identified and removed. The references of the previously selected articles were explored to identify additional literature.
Concluding the literature search, all articles that remained in the bibliography after the filtering were read in detail, focusing on understanding what models were applied to infant cry classification and how these models work. The relevance of the articles was confirmed and supervised-learning classification models were extracted.

2.2. Classification Model Review

For conducting the model review, a common framework for comparing and rating the classification models was defined. Criteria for classification model quality were identified and weightings were provided. Table 1 summarises the rating scheme that was defined. All criteria are rated by an ordinal value between 0 (lowest rating) and 4 (highest rating). Values that are of metric nature were categorised into these five ratings. Rating criteria are described in the following.

Table 1. Rating scheme for the systematic classification model review

		Fulfillment	
Aspect	**Importance factor**	**Category**	**Points**
Accuracy (Acc)	2	96%–100%	4
		91%–95%	3
		81%–90%	2
		51%–80%	1
		<50%	0
Degree of overfitting (OFit)	1	<5%	4
		5%–10%	3
		10%–20%	2
		20%–30%	1
		>30%	0
Conformability provided (Conf)	1,5	Cut-off values and feature importance	4
		Only cut-off values provided	3
		Only feature importance provided	2
		Basic conformability	1
		No conformability	0

Accuracy is the most important aspect for rating the classification models. It is defined as the *precision* of the model on the *test dataset* (a sample not used for training the classification models, but only for validating their accuracy; see Section 2.3).

$$\text{Accuracy} = \frac{N_{11}}{N},$$

where N_{11} is the number of infant cries that were classified correctly and N the overall number of cries. The higher the accuracy on the test dataset, the better. To be comparable with the other criteria, the accuracy value was categorised into an ordinal accuracy category *Acc* from 0 to 4. 0 is the worst accuracy category (accuracy below 50%) and 4 is the best (accuracy above 96%).
The *degree of overfitting* describes how much the model is generalisable to classify unknown cries correctly. It is computed as the difference between the accuracy of the model on the test dataset and the training dataset (c.f. Section 2.3):

$$\text{Degree of overfitting} = \text{Accuracy}_{Test} - \text{Accuracy}_{Training.}$$

Small values indicate better generalisability. The degree of overfitting is categorised, too. The *OFit* categories are from 0 (worst category, degree of overfitting higher than 30%) to 4 (best category, degree of overfitting smaller than 5%; negative values are allowed and fall into this best category, too).

Conformability (Conf) depicts how well a classification model can be conformed by experts. Ratings are given in categories from 0 (no conformability) to 4 (highest conformability category). Basic conformability (category 1) requires that experts can understand how the prediction of the model came about (i.e., the way a data item was categorised is transparent). For higher conformability values, the model must provide information about the importance of the exploratory variables (category 2) or about cut-off values describing what value ranges are typical for a predicted group (category 3; here, the knowledge about value ranges was rated higher than the knowledge about which variables were most influencing for model predictions). If feature importance as well as cut-off values are provided, the highest rating (category 4) is given. Higher conformability ratings are better for evaluating the correctness of the classification model. As the rating of conformability might be subjective, this factor was rated by two independent reviewers. In cases where ratings differed between reviewers, the ratings were discussed and one rating acceptable to both reviewers was picked.

The overall rating (R) was computed as the weighted average of the three categorised criteria:

$$R = \frac{2 \cdot Acc + 1 \cdot OFit + 1.5 \cdot Conf}{4.5}.$$

Therefore, the overall rating takes a value between 0 (worst rating) and 4 (best rating). The selection of weights is discussed in Section 4.

After having defined the rating scheme, four software systems were evaluated on their ability to compute the classification models that had been identified during the literature search; two proprietary software systems (IBM SPSS Statistics 20 (IBM, 2013b) and IBM SPSS Modeler 15 (IBM, 2013a)) and two open-source systems (R 3.0.2 (The R Foundation for Statistical Computing, 2014) and RapidMiner 6 (Rapid-I, 2014)). These four systems implemented most of the classification algorithms. Models for which no implementing software was found, had to be excluded from the review.

Classification models were then applied to a reference dataset. The reference dataset contained 468 cry samples from healthy infants as well as infants with various disorders. The reference dataset is described in Section 2.3 in detail.

Classification models had many parameters influencing the quality of the model. To find the best settings for each model, different settings were tested automatically. By this, all relevant combinations of parameter settings were evaluated to find the best parameter setting for each model.

Classification models were rated according to the rating scheme and results were documented. Model ratings were assessed and suggestions about which classification model approach to use for infant cry classification were provided. The interpretation of our findings is provided in Section 4.1.

2.3. Data Selection

Altogether, cry signals were recorded from 69 infants. Thirty-one of these infants were healthy, full-term babies without any pregnancy complications and indication of physical or neurological disorders. Nineteen infants were hearing impaired with a threshold above 60 dB HL. Ten infants had a unilateral cleft lip and palate (UCLP). Three infants suffering from asphyxia, two infants with brain damage and four infants with laryngomalacia were contributed to the dataset. The infants were between 1 and 7 months of age. For all infants in the pathological groups, it was ensured that they did not suffer from any other disease than the one representing the group. All cries were recorded on a Zoom H2n recorder with a sampling rate of 48 kHz, placed about 30 cm from the infants' mouth. The study was approved by the Ethic Review Committee of the Fresenius University of Applied Science.

Depending on the overall duration of an infant's crying episode, 3 to 11 single-cry utterances were extracted from the episode resulting in 468 cry utterances in total. To allow the rating of model accuracy, the dataset was split into a training dataset used for training a classification model and a separate test dataset used only for rating the model accuracy. Thirty per cent of the cry samples from each group were allocated to the test dataset by chance; the remaining 70% of samples formed the training dataset. Table 2 provides an overview about group sizes for the training and test dataset.

Table 2. Number of cry samples per group and dataset

	Healthy	Hearing impaired	UCLP	Asphyxia	Brain damage	Laryngo-malacia	Sum
Training dataset	200	28	21	11	18	43	321
Test dataset	86	13	7	8	6	27	147
Sum	286	41	28	19	24	70	468

2.4. Acoustic Analysis

The cries were analysed acoustically with Praat Version 5.3.39 (Boersma and Weenink, 2013b). Overall, 19 acoustic parameters were computed for each cry. For the fundamental frequency (F0), Praat's autocorrelation algorithm (Boersma, 1993) was used to calculate the median, the interquartile range and the 90th and 10th percentiles as lower and upper bounds of F0, with a setting between 100 Hz and 1000 Hz.The median of the first six formants (F1–F6) as references to frequency ranges with high spectral intensities were computed with the Burg algorithm (Press et al., 2002). The upper frequency was set to 8000 Hz. The microvariability of vocal fold vibration (Jitter and Shimmer values) was analysed with the waveform-matching algorithm (Boersma, 2009). For the intensity, the median, the interquartile range and the 10th and 90th percentiles were computed with Praat's intensity algorithm. With the forward cross-correlation analysis (Boersma and Weenink, 2013a), the mean of the harmonic-to-noise ratio (HNR) and its standard deviation were estimated. The number and the degree of voice breaks were calculated as well as the fraction of unvoiced pitch frames. The duration of each cry utterance was also measured. Cry parameters were aggregated for each cry, that is, they were stationary and no temporal development was captured (stationary analysis is a common approach in infant cry research; c.f. Michelsson et al., 2002; Robb et al., 2007; Rautava et al., 2007; Branco et al., 2007; Goberman and Robb, 2005).

3. RESULTS

3.1. Systematic Literature Search

Overall, 579 articles were found by the initial search. Table 3 summarises how many articles were found in each database.

Table 3. Search result statistics for the different databases

Database	No. of articles in search result set
Web of Knowledge	94
DIMDI	45
Medpilot	138
DBLP	0
IEEE	35
SSG	17
ACM	250
Sum	579

Sixty-four out of 579 articles remained in the bibliography after filtering (31 articles from medical bibliography databases and 33 articles from computer science bibliographies). No additional literature was identified by exploring the bibliography of the 64 relevant articles.
By reading the articles, nine different types of classification models that had been used for infant cry classification were identified. Table 4 summarises the models and lists studies that used them.

Table 4. List of classification model types and the studies in which they were used

Bayes Classifier
Amaro-Camargo and Reyes-Garcia (2007)
Hidden Markov Model
Aucouturier et al. (2011), Lederman et al. (2002), Lederman et al. (2008), Abdulaziz and Ahmad (2010), Honda et al. (2012), Singh et al. (2013)
Linear Discriminant Analysis (LDA)
Fuller (1991)
Support Vector Machine (SVM)
Brahnam et al. (2006), Lu Guanming et al. (2008), Sahak et al. (2010a), Amaro-Camargo and Reyes-Garcia (2007), Sahak et al. (2010b), Sahak et al. (2011), Sahak et al. (2012)
Fuzzy Logic
Cano-Ortiz et al. (2013), Kia et al. (2012), Reyes-Galaviz et al. (2004), Santiago-Sanchez et al. (2009), Barajas and Reyes (2005)
Decision Tree
Amaro-Camargo and Reyes-Garcia (2007), Etz et al. (2012)
K-Nearest Neighbour (KNN)
Cohen and Lavner (2012)
Weighted Rough Set Framework
Own and Abraham (2012)

3.2. Classification Model Review

The nine identified classification models are described in the following.
Artificial neural networks encompass different machine learning approaches following functions of animal brains by simulating information flow through systems of interconnected 'neurons'. For cry analysis, various kinds of neural networks have been applied, for example, multilayer perceptrons, radial basis function networks, self-organising feature maps and others. The *Bayes classifier* is a probabilistic model based on Bayes' theorem describing classes by statistical processes. *Hidden Markov models* simulate Markov processes with hidden states and are widely used to identify patterns in temporal data like acoustic signals. *Linear discriminant analysis* identifies linear functions to separate groups in data. *Support vector machines* work similar to linear

discriminant analysis, except they can be extended for non-linear discrimination between datasets. *Fuzzy logic* associates certainty values to data, about the possibility of the data item belonging to given groups. *Decision trees* cover different algorithms to compute hierarchical decision rules to decide, to which group data items belong. *K-nearest neighbour* associates data to groups by analysing the -most-similar neighbour data items of an item to be classified. *Weighted rough set framework* targets the class imbalance problem (i.e., groups have different sizes) by providing a mathematical model based on lower and upper approximation of groups.

The four identified software systems (IBM SPSS Statistics 20, IBM SPSS Modeler 15, R 3.0.2, RapidMiner 6) implemented most of the classification algorithms. However, three algorithms could not be evaluated in the review.

Classification algorithms based on *fuzzy logic* were neither implemented in one of the given systems, nor was it possible to find any ready-to-use software system providing fuzzy classification. Therefore, fuzzy classification could not be applied to the dataset. The same problem occurred for *weighted rough sets*; no ready-to-use software was found, either.

Hidden Markov models were available as plug-in for one of the given systems. However, for Markov models it became apparent that they were not suited for the given dataset: Markov models are focused on temporal data, whereas the reference dataset is based on stationary parameters. To keep classification models comparable, Markov models were excluded as they would require different data from the other classification models.

Apart from the classification models identified in the infant cry classification literature, the software systems provided four additional classification models, which were included in the review; three additional classification tree approaches (C&R decision tree, CHAID decision tree and Quest decision tree) and logistic regression. Logistic regression models provide linear formulae for classifying items into categorial groups.

Figure 1 summarises all classification models that were applied to the reference dataset in this article. Algorithms in filled boxes have already been used in the literature and algorithms with blank boxes have not yet been used in the literature and were included in the review.

The classification models were applied to the reference dataset and rated according to the rating scheme.

Table 5 provides the results of the classification model and ratings. Ratings are briefly explained in the following.

3.2.1. Decision Tree Ratings

Five different algorithms for training decision trees were applied to the reference dataset. Accuracy on the test dataset varied between 69.32% and 98.64% resulting in accuracy ratings *Acc* from 1 to 4. The highest degree of overfitting (difference between the accuracy on training and test datasets) was 24.80%, the lowest was -1.44% (i.e., the accuracy on the test dataset was higher than on the training dataset). Ratings for overfitting *OFit* were between 1 and 4. Conformability of all decision trees was rated with 4 ('Cut-off values and feature importance provided') because decision rules within the trees allow identifying characteristic properties of each cry group.

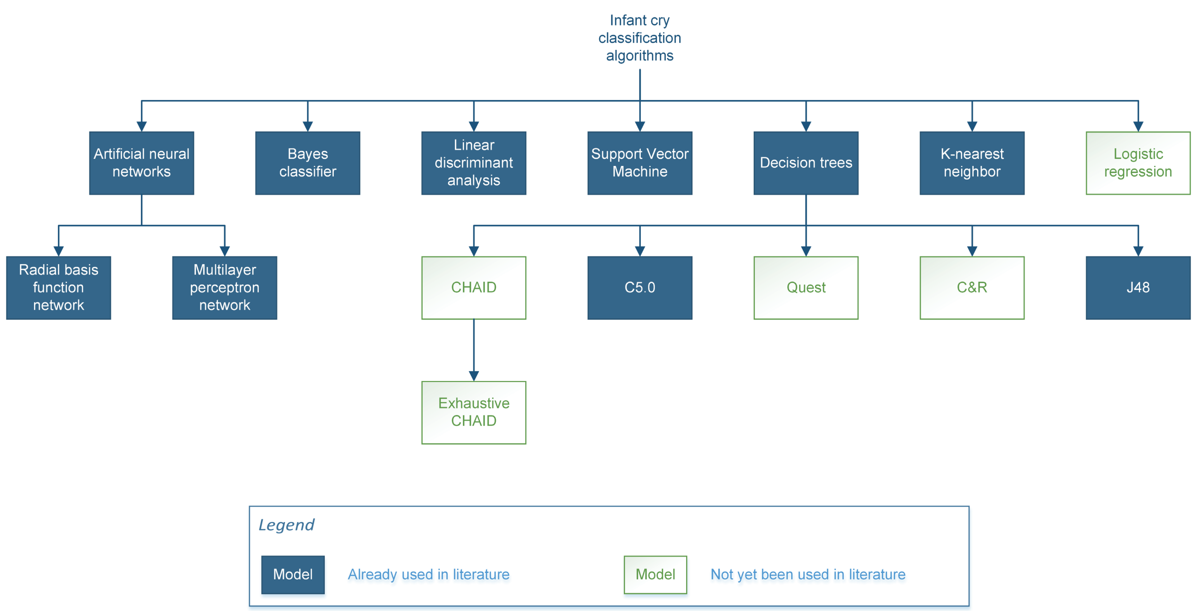

Fig 1. Overview of classification model algorithms that have been included in the model review

Table 5. Rating results for the classification models

Classification model	No. of computed models	Accuracy training dataset	Accuracy test dataset		Degree of overfitting		Conformability		Rating
		% correct	% correct	Rating Acc	Value	Rating OFit	Category	Rating Conf	
Decision trees									
C & R decision tree	144	98,10%	73,30%	1	24,80%	1	Cut-off values and feature importance	4	2
Quest decision tree	48	89,13%	69,32%	1	19,81%	2	Cut-off values and feature importance	4	2,22
Chaid decision tree	384	99,46%	77,84%	1	21,62%	1	Cut-off values and feature importance	4	2
C5 decision tree	96	97,20%	98,64%	4	-1,44%	4	Cut-off values and feature importance	4	4
J48 decision tree	64	97,63%	82,93%	2	14,70%	2	Cut-off values and feature importance	4	2,67
K-nearest neighbour (KNN)	128	95,92%	91,48%	3	4,44%	4	Basic	1	2,56
Bayes classifier	128	91,03%	80,68%	2	10,35%	2	Cut-off values	3	2,33
Linear discriminant analysis	160	90,49%	79,55%	1	10,94%	2	Feature importance	2	1,56
Logistic regression	576	95,11%	69,89%	1	25,22%	1	Feature importance	2	1,33
Support vector machine (SVM)	384	52,17%	53,98%	1	-1,81%	4	Feature importance	2	2
Artificial neural networks									
Neural network: multilayer perceptron	192	96,30%	79,50%	1	16,80%	2	Feature importance	2	1,56
Neural network: radial basis function	48	78,70%	73,90%	1	4,80%	4	Feature importance	2	2

The overall rating ($R = \frac{2 \cdot Acc + OFit + 1.5 \cdot Conf}{\sum \text{Importance factors}}$) was between 2.00 and 4.0 for the decision trees.

3.2.2. K-nearest Neighbour

The best k-nearest neighbour model reached an accuracy value of 91.48% on the test dataset resulting in a rating of 3 for *Acc*. A degree of overfitting at 4.44% led to an *OFit* rating of 4. Conformability of k-nearest neighbour models were rated with the basic degree (*Conf*=1). No characteristic cut-off values could be extracted from the model representation. In addition, feature importance was limited to the top three most important features. Further information about the importance of the remaining acoustic properties of the cry signal is missing. Overall, the k-nearest neighbour model was rated with 2.56.

3.2.3. Bayes Classifier

The Bayes classifier had an accuracy of 80.68% on the test dataset (*Acc*=2) and an overfitting of 10.35% (*OFit*=2). Cut-off values were recognisable by exploring the probability statistics of the Bayes classifier results (*Conf*=3). Summarising, the Bayes classifier got a rating of 2.33.

3.2.4. Linear Discriminant Analysis

Linear discriminant analysis achieved 79.55% accuracy on the test dataset resulting in a rating *Acc* of 1. With a degree of overfitting at , the model reached a rating *OFit* of 2. Cut-off values for each acoustic feature were not recognisable. However, information about the feature importance was provided, resulting in a rating *Conf* of 2. The overall rating for linear discriminant analysis was 1.56.

3.2.5. Logistic Regression

The best logistic regression model reached a rating *Acc* of 1 with 69.89% accuracy on the test dataset; the overfitting of 25.22% was rated with *OFit*=1. Similar to linear discriminant analysis models, only the feature importance was recognisable, and therefore, *Conf* was rated with 2. Overall, logistic regression got a rating of 1.33.

3.2.6. Support Vector Machine

The best support vector machine model only reached an accuracy of 53.98% on the test dataset resulting in a rating *Acc* of 1. However, accuracy on the test dataset was even higher than on the training dataset, providing a high rating *OFit*=4 for the degree of overfitting. As for discriminant analysis, no cut-off values were provided; this resulted in a rating *Conf* of 2 for providing only feature importance information. Aggregating the single rating values resulted in an overall rating of 2.00 for the support vector machine.

3.2.7. Artificial Neural Networks

Two types of artificial neural networks were trained, multilayer perceptron based network and radial basis function based network. Accuracy values for these two types of models were 79.50% and 73.90%, respectively, on the test dataset; both values resulted in a rating of 1 for . A degree of overfitting at 16.80% resulted in a rating of 2 for *OFit* for the multilayer preceptron network; *OFit* for the radial basis network was rated with 4 (4.80%). Both types of artificial neural networks only provide information about the feature importance (*Conf*=2). No information about what characteristics of the cry signal lead to a prediction was given. Overall, multilayer preceptron networks were rated with 1.56; radial basis function networks were rated with 2.00 in summary.

4. DISCUSSION

The discussion section is split into two parts: the interpretation of the review results (which is part of the review process) is presented first and threats to validity are discussed thereafter. The article ends with a conclusion and outlook on future work.

4.1. Interpretation of Review Results

The systematic literature search identified much research that explored supervised-learning models for infant cry classification. Many different approaches have been used so far, but no systematic comparison of models has been conducted.

On the training dataset, most classification models achieved high accuracy values confirming the promising results reported in the literature. However, when validating the models on an independent test dataset, accuracy often dropped significantly. On the test dataset, only two models have accuracy values beyond 90% (C5.0 decision trees with 98.64% and K-nearest neighbour with 91.48%). The remaining models seem to have serious problems with correctly predicting an infant's state of health by her cry.

Additionally, the gap between high accuracy on the training dataset and much lower accuracy on an independent test dataset indicates another problem of classification models: model overfitting. Overfitting occurs when models are too specialised on the training dataset. It leads to a lack of abstraction and therefore prevents correct classification of data that is slightly different from the training dataset.

Analysis of model validity should not depend on statistical values only. Therefore, understanding how models predict data and to verify the correctness of this prediction by experts, is very important. Here, several models suffered from too complex algorithms and too little insight into the classification process. Here, decision tree approaches provided the best conformability as they allow to verify the decision rules of the trees visually. However, trees should not grow too large as they then can get very complicated to understand as well.

Overall, suitability of classification models for infant cry analysis varied between high suitability and low suitability. Interestingly, classification models that have not been explored in infant cry research very well (e.g., several decision tree approaches or k-nearest neighbour) performed better than modelling approaches that have often been used in infant cry classification (e.g., neural networks). Based on this observation, our major recommendation derived from the model review would be to give those 'exotic' classification models a try and explore their suitability for infant cry classification in more detail in future research.

4.2. Threats to Validity

The validity of systematic review processes may be influenced by various factors. In this systematic classification review, the following potential factors were identified and addressed to improve validity as much as possible.

First, the rating framework might be biased to favour certain models. For this reason, a short literature search was performed prior to the review, in order to identify rating factors used in other classification model reviews (classifying various data). Parameters were selected and adapted to fit the classification of infant cries. In addition, the weighting of the single rating factors is arbitrary. Here, the selection followed the authors' view on what is important for infant cry classification. As the most important factor, the accuracy of the model on an independent test dataset was chosen. With a slightly lower priority — but still very important — the conformability was seen. In the authors' opinion, it is not sufficient to rely on statistical accuracy values but to be able to validate model ratings by experts; this requires high transparency of the rating process. As the least important aspect, the degree of overfitting was chosen as it is not quite as important as the first two factors.

The reference dataset was constructed to cover healthy cries as well as multiple pathologies in order to identify the models' abilities to discriminate among many groups. For some pathologies, it is difficult to recruit many subjects resulting in different group sizes in the dataset. As a threat to validity, this might introduce bias when training classification models. However, for infant cry classification used as an early indicator of an infant's state of health, it is more important to be able to discriminate between healthy and non-healthy cries. Here, the reference dataset contains about 60% healthy cry samples and 40% non-healthy ones. Classification models were rated on their overall accuracy and not on their ability to predict certain groups for this reason.

Quality (especially the accuracy) of models might vary, depending on the parameter selection for the classification algorithms. For this reason, multiple combinations of reasonable parameter values were selected and models were trained with each parameter setting, automatically. By this, the parameters providing the model with the highest quality were selected for each classification approach to ensure that the best result was achieved.

4.3. Conclusion and Future Work

Summarising, this article presents a systematic review of classification models for infant cry classification. In a systematic literature search covering seven well-known databases of medical and computer science publications, 579 articles were analysed to extract 12 different classification model approaches that have already been used in the literature for infant cry classification. Nine of those approaches plus three additional approaches that have not yet been used for infant cry classification were enabled and applied to a reference dataset containing 468 cry samples from 69 infants grouped into six different states of health. Models were rated systematically according to a predefined rating framework. Results indicate that many models have issues with accuracy and conformability (i.e., if the validity of models can be confirmed by experts by exploring the models). However, some of the models provide promising results in infant cry classification.

Some of the models that have not yet been explored very well in the infant cry classification literature achieved better results than approaches that have often been used for classification. Our recommendation for future research is, to give those approaches a try and explore their suitability for infant cry classification in more detail. Concluding, the findings of our research show that classification models exist that perform well in classifying healthy and pathological infant cries and therefore could be used to develop a screening instrument based on the acoustic characteristics of infant cries for identifying various kinds of pathological development, early. Such screening instruments may in future be helpful for various health professions concerned with infants like nurses, midwifes, therapists or pediatricians.

References

Abdulaziz, Y., & Ahmad, S. (2010). An accurate infant cry classification system based on continuous Hidden Markov Model. In Information Technology (ITSim), 2010 International Symposium in (Vol. 3, pp. 1648–1652).

ACM. (2014). ACM Digital Library. Retrieved from http://dl.acm.org/dl.cfm

Amaro-Camargo, E., & Reyes-Garcia, C. A. (2007). Applying statistical vectors of acoustic characteristics for the automatic classification of infant cry. Advances in Pattern Recognition, 4681, 1078–1085.

Arch-Tirado, E., Mandujano, M., Garcia-Torices, L., Martinez-Cruz, C. F., Reyes-García, C. A., & Taboada-Picazo, V. (2004). Cry analysis of hypoacoustic children and normal hearing children. Cirugia y Cirujanos, 72(4), 271–276.

Aucouturier, J.-J., Nonaka, Y., Katahira, K., & Okanoya, K. (2011). Segmentation of expiratory and inspiratory sounds in baby cry audio recordings using hidden Markov models. Journal of the Acoustical Society of America, 130(5), 2969–2977.

Barajas, S. E., & Reyes, C. A. (2005). Your fuzzy relational neural network parameters optimization with a genetic algorithm. FUZZ-IEEE 2005: Proceedings of the IEEE International Conference on Fuzzy Systems: BIGGEST LITTLE CONFERENCE IN THE WORLD, 684–689.

Barajas-Montiel, S., & Reyes-Garcia, C. (2005). Identifying Pain and Hunger in Infant Cry with Classifiers Ensembles. In Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on (Vol. 2, pp. 770–775).

2012. Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on.

Blinick, G., Tavolga, W. N., & Antopol, W. (1971). Variations in birth cries of newborn infants from narcotic-addicted and normal mothers. American Journal of Obstetrics and Gynecology, 110(7), 948–958.

Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. In Proceedings of the Institute of Phonetic Sciences, Amsterdam (Vol. 17, pp. 97–110).

Boersma, P. (2009). Should Jitter Be Measured by Peak Picking or by Waveform Matching? Folia Phoniatrica et Logopaedica, 61(5), 305–308.

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer. Version 5.3.39.

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer: Manual. Retrieved from http://www.fon.hum.uva.nl/praat/manual/Intro.html

Brahnam, S., Chuang, C. F., Shih, F. Y., & Slack. (2006). Machine recognition and representation of neonatal facial displays of acute pain. Artificial Intelligence in Medicine, 36(3), 211–222.

Branco, A., Fekete, S. M., Rugolo, L. M., & Rehder, M. I. (2007). The newborn pain cry: descriptive acoustic spectrographic analysis. International Journal of Pediatric Otorhinolaryngology, 71(4), 539–546.

Cano, S., Suaste, I., Escobedo, D., Reyes-Garcia, C. A., & Ekkel, T. (2006). A combined classifier of cry units with new acoustic attributes. Advances in Pattern Recognition, 4225, 416–425.

Cano-Ortiz, S. D., Reyes-Garcia, C. A., Reyes-Galaviz, O. F., Escobedo-Beceiro, D. I., & Cano-Otero, J. D. (2013). Emergence of a New Alternative on Cry Analysis: The Fuzzy Approach. 5TH LATIN AMERICAN CONGRESS ON BIOMEDICAL ENGINEERING (CLAIB 2011): SUSTAINABLE TECHNOLOGIES FOR THE HEALTH OF ALL, PTS 1 AND 2, 33(1-2), 846–849.

Chollet, G., Esposito, A., Faundez-Zanuy, M., & Marinaro, M. (Eds.). (2005). Berlin and Heidelberg: Springer-Verlag.

Cohen, R., & Lavner, Y. (2012). Infant cry analysis and detection. In Electrical Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of (pp. 1–5).

Corwin, M. J., Lester, B. M., Sepkoski, C., McLaughlin, S., Kayne, H., & Golub, H. L. (1992). Effects of in utero cocaine exposure on newborn acoustical cry characteristics. Pediatrics, 89(6 Pt 2), 1199–1203.

DBLP. (2014). Digital Bibliography & Library Project. Retrieved from http://www.dblp.org/

Deutsche Zentralbibliothek für Medizin. (2014). MEDPILOT. Retrieved from http://www.medpilot.de/

DIMDI. (2014). DIMDI Medizinwissen. Retrieved from http://www.dimdi.de/static/de/index.html

Esposito, G., Nakazawa, J., Venuti, P., & Bornstein, M. H. (2013). Componential deconstruction of infant distress vocalizations via tree-based models: A study of cry in autism spectrum disorder and typical development. Research in Developmental Disabilities, 34(9), 2717–2724.

Etz, T., Reetz, H., & Wegener, C. (2012). A classification model for infant cries with hearing impairment and unilateral cleft lip and palate. Folia Phoniatrica et Logopaedica, 64(5), 254–261.

Fisichelli, V. R., & Karelitz, S. (1966). Frequency spectra of the cries of normal infants and those with Down's syndrome. Psychonomic Science, 6, 195–196.

Fort, A., & Manfredi, C. (1998). Acoustic analysis of newborn infant cry signals. Medical Engineering & Physics, 20(6), 432–442.

Fuller, B. F. (1991). Acoustic discrimination of three types of infant cries. Nursing Research, 40(3), 156–160.

Galaviz, Orion Fausto Reyes, & García, Carlos Alberto Reyes (2005). Infant cry classification to identify hypo acoustics and asphyxia comparing an evolutionary-neural system with a neural network system. In: MICAI'05, Proceedings of the 4th Mexican international conference on Advances in Artificial Intelligence (pp. 949–958). Springer-Verlag.

Garcia, J., & Reyes Garcia, C. (2003). Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks. In Neural Networks, 2003. Proceedings of the International Joint Conference on (Vol. 4, pp. 3140-3145 vol. 4).

Garcia, J. O., & Reyes Garcia, Carlos A. (2003). Medical Applications - Acoustic Features Analysis for Recognition of Normal and Hipoacustic Infant Cry Based on Neural Networks. Lecture Notes in Computer Science, 2687, 615–622.

GBV. (2014). Online Contents - SSG Mathematik und Informatik. Retrieved from http://gso.gbv.de/DB=2.77/LNG=DU/

Goberman, A. M., & Robb, M. P. (2005). Acoustic characteristics of crying in infantile laryngomalacia. Logopedics, Phoniatrics, Vocology, 30(2), 79–84.

Golub, H. L., & Corwin, M. J. (1982). Infant cry: a clue to diagnosis. Pediatrics, 69(2), 197–201.

Hariharan, M., Saraswathy, J., Sindhu, R., Khairunizam, W., & Yaacob, S. (2012). Infant cry classification to identify asphyxia using time-frequency analysis and radial basis neural networks. Expert Systems with Applications, 39(10), 9515–9523.

Hariharan, M., Saraswathy, J., Sindhu, R., Khairunizam, W., & Yaacob, S. (2012). Infant cry classification to identify asphyxia using time-frequency analysis and radial basis neural networks. Expert Systems with Applications, 39(10), 9515–9523.

Hariharan, M., Sindhu, R., & Yaacob, S. (2012). Normal and hypoacoustic infant cry signal classification using time-frequency analysis and general regression neural network. Computer Methods and Programs in Biomedicine, 108(2), 559–569.

Hariharan, M., Yaacob, S., & Awang, S. A. (2011). Pathological infant cry analysis using wavelet packet transform and probabilistic neural network. Expert Systems with Applications, 38(12), 15377–15382.

Honda, K., Kitahara, K., Matsunaga, S., Yamashita, M., & Shinohara, K. (2012). Emotion classification of infant cries with consideration for local and global features. In Signal Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific (pp. 1–4).

IBM. (2013). SPSS Modeler. Version 15.

IBM. (2013). SPSS Statistics.

IEEE. (2014). IEEE Xplore Digital Library. Retrieved from http://ieeexplore.ieee.org/

Karelitz, S., & Fisichelli, V. R. (1962). The cry thresholds of normal infants and those with brain damage. Disability and Rehabilitation, 61(5), 679–684.

Kheddache, Y., & Tadj, C. (2012). Newborn's pathological cry identification system. In Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on (pp. 1024–1029).

Kia, M., Kia, S., Davoudi, N., & Biniazan, R. (2012). A detection system of infant cry using fuzzy classification including dialing alarm calls function. In Innovative Computing Technology (INTECH), 2012 Second International Conference on (pp. 224–229).

Lederman, D., Cohen, A., Zmora, E., Wermke, K., Hauschildt, S., & Stellzig-Eisenhauer, A. (2002). On the use of hidden Markov models in infants' cry classification. 22ND CONVENTION OF ELECTRICAL AND ELECTRONICS ENGINEERS IN ISRAEL, PROCEEDINGS, 350–352.

Lederman, D., Zmora, E., Hauschildt, S., Stellzig‐Eisenhauer, A., & Wermke, K. (2008). Classification of cries of infants with cleft-palate using parallel hidden Markov models. Medical & Biological Engineering & Computing, 46(10), 965–975.

Lester, B. M., Tronick, E. Z., LaGasse, L. L., Seifer, R., Bauer, C. R., Shankaran, S., . . . Maza, P. L. (2002). The maternal lifestyle study: effects of substance exposure during pregnancy on neurodevelopmental outcome in 1-month-old infants. Pediatrics, 110(6), 1182–1192.

Lind, J., Vuorenkoski, V., & Rosberg, G. (1970). Spectrographic analysis of vocal response to pain stimuli in infants with Down's syndrome. Developmental Medicine and Child Neurology, 12(4), 478–486.

Lind, J., Wasz-Höckert, O., Rosberg, G., Theorell, K., Valanne, E. H., Partanen, T. J., & Vuorenkoski, V. (1967). Sound spectrography in pediatric diagnosis. Acta Paediatrica Scandinavica, 177(Suppl. s177), 113–119.

Lu Guanming, Li Xiaonan, & Li Haibo. (2008). Research on Recognition for Facial Expression of Pain in Neonates. Acta Optica Sinica, 28(11), 2109–2114.

Michelsson, K., Eklund, K., Leppänen, P., & Lyytinen, H. (2002). Cry Characteristics of 172 Healthy 1- to 7-Day-Old Infants. Folia Phoniatrica et Logopaedica, 54(4), 190–200.

Michelsson, K., Sirvio, P., & Wasz-Hockert, O. (1977). Sound spectrographic cry analysis of infants with bacterial meningitis. Developmental Medicine and Child Neurology, 19(3), 309–315.

Mohd Ali, M., Mansor, W., Lee, Y. K., & Zabidi, A. (2012). Asphyxiated infant cry classification using Simulink model. In Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on (pp. 491–494).

Möller, S., & Schönweiler, R. (1999). Analysis of infant cries for the early detection of hearing impairment. Speech Communication, 28(3), 175–193.

Nugent, J. K., Lester, B. M., Greene, S. M., Wieczorek-Deering, D., & O'Mahony, P. (1996). The effects of maternal alcohol consumption and cigarette smoking during pregnancy on acoustic cry analysis. Child Development, 67(4), 1806–1815.

Orozco, J., & Reyes-Garcia, C. A. (2003). Implementation and analysis of training algorithms for the classification of infant cry with feed-forward neural networks. 2003 IEEE INTERNATIONAL SYMPOSIUM ON INTELLIGENT SIGNAL PROCESSING, PROCEEDINGS: FROM CLASSICAL MEASUREMENT TO COMPUTING WITH PERCEPTIONS, 271–276.

Orozco-Garcia, J., Reyes-Garcia, C. A., & Reyes-Garcia, C. (2003). A Study on the Recognition of Patterns of Infant Cry for the Identification of Deafness in Just Born Babies with Neural Networks. Lecture Notes in Computer Science, 2905, 342–349.

Ortiz, S. D., Beceiro, D. I., & Ekkel, T. (2004). A radial basis function network oriented for infant cry classification. Advances in Pattern Recognition, 3287, 374–380.

Ortiz, Sergio D. Cano, Escobedo Beceiro, Daniel L., Ekkel, T., & Ortiz, S. D. (2004). A Radial Basis Function Network Oriented for Infant Cry Classification. Lecture Notes in Computer Science, 3287, 374–380.

Own, H. S., & Abraham, A. (2012). A new weighted rough set framework based classification for Egyptian NeoNatal Jaundice. Appl. Soft Comput., 12(3), 999–1005.

Petroni, M., Malowany, M., Johnston, C. C., & Stevens, B. J. (1995). A comparison of neural network architectures for the classification of three types of infant cry vocalizations. In Engineering in Medicine and Biology Society, 1995., IEEE 17th Annual Conference (Vol. 1, pp. 821–822 vol. 1).

Petroni, M., Malowany, M., Johnston, C. C., & Stevens, B. J. (1995). On the use of artificial neural networks (ANNs) for the classification of three types of infant cries. In Communications, Computers, and Signal Processing, 1995. Proceedings., IEEE Pacific Rim Conference on (pp. 501–504).

Poel, M., & Ekkel, T. (2006). Analyzing infant cries using a committee of neural networks in order to detect hypoxia related disorder. International Journal on Artificial Intelligence Tools, 15(3), 397–410.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (2002). Numerical recipes in C: The art of scientific computing (2nd ed.). Cambridge: Cambridge University Press.

Rapid-I. (2014). RapidMiner Studio 6. Retrieved from http://rapidminer.com/products-2/rapidminer-studio/

Rautava, L., Lempinen, A., Ojala, S., Parkkola, R., Rikalainen, H., Lapinleimu, H., . . . Lehtonen, L. (2007). Acoustic quality of cry in very-low-birth-weight infants at the age of 1 1/2 years. Early Human Development, 83(1), 5–12.

Reyes Galaviz, Orion Fausto, & Reyes Garcia, Carlos Alberto. (2004). Intelligent Interfaces and Speech Processing - Infant Cry Classification to Identify Hypoacoustics and Asphyxia with Neural Networks. Lecture Notes in Computer Science, 2972, 69–78.

Reyes Galaviz, Orion Fausto, & Reyes Garcia, Carlos Alberto. (2005). Bioinformatics and Medical Applications - Infant Cry Classification to Identify Hypo Acoustics and Asphyxia Comparing an Evolutionary-Neural System with a Neural Network System. Lecture Notes in Computer Science, 3789, 949–958.

Reyes-Galaviz, O. F., Arch Tirado, E., Reyes-Garcia, C. A., Reyes-Galaviz, O., & Reyes-Garcia, C. (2004). Elderly and Disabled People: Therapy and Care - Classification of Infant Crying to Identify Pathologies in Recently Born Babies with ANFIS. Lecture Notes in Computer Science, 3118, 408–415.

Reyes-Galaviz, O., Cano-Ortiz, S., & Reyes-Garcia, C. (2008). Validation of the Cry Unit as Primary Element for Cry Analysis Using an Evolutionary-Neural Approach. In Computer Science, 2008. ENC '08. Mexican International Conference on (pp. 261–267).

Reyes-Galaviz, O. F., Verduzco, A., Arch-Tirado, E., & Reyes-García, C. A. (2005). Analysis of an infant cry recognizer for the early identification of pathologies: Nonlinear Speech Modeling and Applications. In G. Chollet, A. Esposito, M. Faundez-Zanuy, & M. Marinaro (Eds.) (pp. 404–409). Berlin and Heidelberg: Springer-Verlag.

Reyes-Garcia, C. A., Reyes-Galaviz, O. F., Cano-Ortiz, S. D., Escobedo-Becerro, D. I., Zatarain, R., & Barron-Estrada, L. (2010). Soft Computing Approaches to the Problem of Infant Cry Classification with Diagnostic Purposes. Soft Computing for Recognition Based on Biometrics, 312, 3–18.

Robb, M. P., Crowell, D. H., & Dunn-Rankin, P. (2007). Cry analysis in infants resuscitated for apnea of infancy. International Journal of Pediatric Otorhinolaryngology, 71(7), 1117–1123.

Rosales-Pérez, A., Reyes-García, C. A., & Gómez-Gil, P. (2011). Genetic fuzzy relational neural network for infant cry classification. In : MCPR'11, Proceedings of the Third Mexican Conference on Pattern Recognition (pp. 288–296). Springer-Verlag.

Sahak, R., Lee, Y., Mansor, W., Yassin, A. I. M., & Zabidi, A. (2010). Optimized Support Vector Machine for classifying infant cries with asphyxia using Orthogonal Least Square. In Computer Applications and Industrial Electronics (ICCAIE), 2010 International Conference on (pp. 692–696).

Sahak, R., Lee, Y., Mansor, W., Zabidi, A., & Yassin, A. I. M. (2011). Detection of asphyxia from infant cry by linear kernel support vector machine enhanced with features from orthogonal least square. In Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference on (pp. 341–345).

Sahak, R., Mansor, W., Lee, Y. K., Yassin, A. I. M., & Zabidi, A. (2010). Performance of Combined Support Vector Machine and Principal Component Analysis in recognizing infant cry with asphyxia. In 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 6292–6295). IEEE.

Sahak, R., Mansor, W., Khuan, L., Zabidi, A., & Yassin, A. I. M. (2012). Detection of asphyxia from infant cry using support vector machine and multilayer perceptron integrated with Orthogonal Least Square. In Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on (pp. 906–909).

Santiago-Sanchez, K., Reyes-Garcia, C. A., & Gomez-Gil, P. (2009). Type-2 Fuzzy Sets Applied to Pattern Matching for the Classification of Cries of Infants under Neurological Risk. Advances in Pattern Recognition, 5754, 201–210.

Saraswathy, J., Hariharan, M., Vijean, V., Yaacob, S., & Khairunizam, W. (2012). Performance comparison of Daubechies wavelet family in infant cry classification. In Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on (pp. 451–455).

Schonweiler, R., Kaese, S., Moller, S., Rinscheid, A., & Ptok, M. (1996). Neuronal networks and self-organizing maps: New computer techniques in the acoustic evaluation of the infant cry. International Journal of Pediatric Otorhinolaryngology, 38(1), 1–11.

Schönweiler, R., Kaese, S., Möller, S., Rinscheid, A., & Ptok, M. (1996). Neuronal networks and self-organizing maps: new computer techniques in the acoustic evaluation of the infant cry. International Journal of Pediatric Otorhinolaryngology, 38(1), 1–11.

Sheinkopf, S. J., Iverson, J. M., Rinaldi, M. L., & Lester, B. M. (2012). Atypical cry acoustics in 6-month-old infants at risk for autism spectrum disorder. Autism Research, 5(5), 331–339.

Singh, A. K., Mukhopadhyay, J., & Rao, K. (2013). Classification of infant cries using epoch and spectral features. In Communications (NCC), 2013 National Conference on (pp. 1–5).

Sirviö, P., & Michelsson, K. (1976). Sound-spectrographic cry analysis of normal and abnormal newborn infants: A review and a recommendation for standardization of the cry characteristics. Folia Phoniatrica, 28(3), 161–173.

Suaste-Rivas, I., Diaz-Méndez, A., Reyes-Garcia, C. A., Reyes-Galaviz, O. F., Reyes-Garcia, C., & Reyes-Galaviz, O. (2006). IV Dialogue - Hybrid Neural Network Design and Implementation on FPGA for Infant Cry Recognition. Lecture Notes in Computer Science, 4188, 703–710.

Suaste-Rivas, I., Reyes-Galaviz, O. F., Diaz-Mendez, A., & Reyes-Garcia, C. A. (2004). A fuzzy relational neural network for pattern classification. Advances in Pattern Recognition, 3287, 358–365.

Suaste-Rivas, I., Reyes-Galviz, O. F., Diaz-Mendez, A., & Reyes-Garcia, C. A. (2004). Implementation of a linguistic fuzzy relational neural network for detecting pathologies by infant cry recognition. Advances in Pattern Recognition, 3315, 953–962.

The R Foundation for Statistical Computing. (2014). The R Project for Statistical Computing. Retrieved from http://www.r-project.org/

Thoden, C. J., & Michelsson, K. (1979). Sound spectrographic cry analysis in Krabbe's Disease. Developmental Medicine and Child Neurology, 21(3), 400–401.

Thomson Reuters. (2014). Web of Science. Retrieved from http://www.isiknowledge.com/WOS

Várallyay, György Jr. (2007). The melody of crying. International Journal of Pediatric Otorhinolaryngology, 71(11), 1699–1708.

Verduzco-Mendoza, A., Arch-Tirado, E., Reyes-Garcia, C. A., Leybon-Ibarra, J., & Licona-Bonilla, J. (2012). Spectrographic cry analysis in newborns with profound hearing loss and perinatal high-risk newborns. Cirugia y Cirujanos, 80(1), 3–10.

Vuorenkoski, V., Lind, J., & Partanen, T. J. (1966). Spectrographic analysis of cries from children with maladie du cri du chat. Annales Paediatriae Fenniae, 12(3), 174–180.

Wasz-Höckert, O., Lind, J., Vuorenkoski, V., Partanen, T. J., & Valanne, E. H. (1968). The infant cry: A spectrographic and auditory analysis (1st ed.). Clinics in developmental medicine: Vol. 29: Cambridge University Press.

Zabidi, A., Khuan, L. Y., Mansor, W., Yassin, I. M., & Sahak, R. (2010). Classification of Infant Cries with Asphyxia Using Multilayer Perceptron Neural Network. 2010 SECOND INTERNATIONAL CONFERENCE ON COMPUTER ENGINEERING AND APPLICATIONS: ICCEA 2010, PROCEEDINGS, VOL. 1, 204–208.

Zabidi, A., Khuan, L., Mansor, W., Yassin, I. M., & Sahak, R. (2011). Binary particle swarm optimization for feature selection in detection of infants with hypothyroidism. In Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE (pp. 2772–2775).

Zabidi, A., Mansor, W., Lee Yoot Khuan, Yassin, I., & Sahak, R. (2010). Discrete Mutative Particle Swarm Optimisation of MFCC computation for classifying hypothyroidal infant cry. In Computer Applications and Industrial Electronics (ICCAIE), 2010 International Conference on (pp. 588–592).

Zabidi, A., Mansor, W., Yoot Khuan Lee, Yassin, I., & Sahak, R. (2011). Binary Particle Swarm Optimization for selection of features in the recognition of infants cries with asphyxia. In Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on (pp. 272–276).

Zabidi, A., Mansor, W., Lee Yoot Khuan, Yassin, I., & Sahak, R. (2011). Three-dimensional particle swam optimisation of Mel Frequency Cepstrum Coefficient computation and Multilayer Perceptron neural network for classifying asphyxiated infant cry. In Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference on (pp. 290–293).