Infant Cry Classification Using Genetic Selection of a Fuzzy Model

Alejandro Rosales-Pérez[1], Carlos A. Reyes-García[1], Jesus A. Gonzalez[1], and Emilio Arch-Tirado[2]

[1] National Institute of Astrophysics, Optics and Electronics (INAOE)
Computer Science Department
Tonantzintla, Puebla, Mexico
{arosales,kargaxxi,jagonzalez}@ccc.inaoep.mx

[2] National Institute of Rehabilitation (INR). Laboratory of Bioacoustics
Mexico City, Mexico
earch@inr.gob.mx

Abstract. In the last years, infant cry recognition has been of particular interest because it contains useful information to determine if the infant is hungry, has pain, or a particular disease. Several studies have been performed in order to differentiate between these kinds of cries. In this work, we propose to use Genetic Selection of a Fuzzy Model (GSFM) for classification of infant cry. GSFM selects a combination of feature selection methods, type of fuzzy processing, learning algorithm, and its associated parameters that best fit to the data. The experiments demonstrate the feasibility of this technique in the classification task. Our experimental results reach up to 99.42% accuracy.

Keywords: Infant Cry Classification, Model Selection, Genetic Algorithms.

1 Introduction

Infant cry is the only initial communication way babies have in their earlier stages of life. By crying, the baby can express either the cause of the crying (pain, hunger, etc.) or the presence of physical abnormalities (pathology or disease). Generally, most parents learn to distinguish causes of crying and interpret the baby's need, but it is not easy to distinguish normal from pathological cry, and even less to distinguish one pathology from another. It has been shown that the acoustic characteristics of the sound of crying are influenced by physical or psychological aspects of the infant as well as external stimuli [11]. In this sense, the crying wave contains useful information to distinguish different states of the infant.

Several studies have been performed around infant cry recognition. In signal processing Mel Frequency Cepstral Coefficients (MFCC) and Linear Prediction Coefficient (LPC) techniques have been applied to extract features to represent the audio signals. Several pattern recognition techniques have been used for the classification task. Among these techniques, artificial neural networks have been

L. Alvarez et al. (Eds.): CIARP 2012, LNCS 7441, pp. 212–219, 2012.

one of the most widely used [6,9,14,16]. It has also been explored the use of support vector machines [3,19], hidden Markov models [12,13], as well as several hybrid approaches that combine fuzzy logic with neural networks [15,20,21,22], fuzzy logic with support vector machines [2] or evolutionary strategies with neural networks [8]. These works have reported promising results in infant cry recognition. However, many efforts had to be devoted in the manual design of the classifiers with the intend to determine the set of adequate parameters for each technique to get the right classification of infant cry. In order to avoid many drawbacks, alternative approaches that combine genetic algorithms with fuzzy logic and neural networks have been proposed [1,17], but most of the works only determine the parameters for a specific learning algorithm. In this work we propose to explore the use of Genetic Selection of a Fuzzy Model (GSFM) for infant cry classification. GSFM was recently proposed and it was applied to acute leukemia classification [18]. A genetic algorithm is used in GSFM for selecting the right combination of a feature selection method, the type of fuzzy processing, a learning algorithm, and their associated parameters that better fit to a data set.

The rest of the paper is organized as follows: in section 2 we describe the data set used in our work. In section 3 our classification approach is described. Next, in section 4 the experiments and results are shown. Finally, in section 5 the conclusions and future work are presented.

2 Data Set Description

The infant cry samples were collected directly by specialized physicians. The samples were labeled in the moment of their recording. Labels contain information about the cause of the cry or the pathology presented.

Recordings were divided in segments of one second, each segment is then taken as an individual sample of cry. Samples were divided in frames of 50 milliseconds. The MFCC technique was applied to the samples and 16 coefficients were extracted from each frame, getting vectors with 304 coefficients. The Pratt [5] tool was used to extract the coefficients.

The infant cry corpus has 340 samples of cries of asphyxia, 192 for pain, 350 for hunger, 879 cries of babies who are deaf and 157 of normal cries. Pain and hunger cries come from normal babies, so they are also part of the normal cries collection. This corpus was used to build different binary data sets: asphyxia vs normal and hunger, deaf vs normal and hunger and hunger vs pain. Table 1 shows the different data sets and the number of samples of each case. These data sets were used in our experiments.

3 Classification Approach

A genetic algorithm, proposed by Holland [10] in the 70s, is a heuristic search technique which is inspired in the Darwin's evolutionary theory to solve problems using computational models. The genetic algorithm is based on the idea of the survival of the individual's fitness and, reproduction strategies, where

Table 1. Description of infant cry data sets

Data set	No. Samples	Samples by class
Asphyxia vs Normal and Hungry	847	Asphyxia: 340 Normal and Hungry: 507
Deaf vs Normal and Hungry	1386	Deaf: 879 Normal and Hungry: 507
Hungry vs Pain	542	Hungry:350 Pain:192

stronger individuals have a higher chance to create offsprings, and consequently are considered in the evolution process. Generally, a genetic algorithm has five basic components: an encoding scheme, in a form of chromosomes or individuals, that represents the potential solutions to the problem, a form to create potential initial solutions, a fitness function to measure how close a chromosome is to the desired solution, selection operations and reproduction operators [7].

In this work we use GSFM for infant cry classification. For each data set, described in section 2, we applied GSFM to select a model. The obtained model was trained with a training data set, and this model was used to predict the testing set. Next, the GSFM technique is described.

3.1 Genetic Selection of a Fuzzy Model

The process of the construction of a fuzzy model is shown in Fig. 1. A labeled data set is the input. Given that each sample is described by a set of N features, and that N is usually large, the first step is to reduce the dimensionality of the data set. This task is done by applying a feature selection method. Then, the subset of selected features is converted into fuzzy values, which is the fuzzification step. Next, the parameters of fuzzy membership are fitted to reduce the overlapping degree. Finally, with the fuzzy features a fuzzy classifier is built. Given a pool of feature selection methods, fuzzy processing and learning algorithms, GSFM selects the combination of them that minimizes the error.

GSFM has the advantage to consider different methods, Table 2 describes each of methods considered by GSFM. Nevertheless, searching across the model space is computationally intractable. In GSFM, each model (combination of different methods) is represented by a chromosome, also known as individual. For each generation in the genetic algorithm, a set of chromosomes is evaluated and these chromosomes are used to generate new models for the next generation. A fitness function is required to asses the models, in this case we considered the balance error rate (BER). However, the error of each model is not known *a priori*, for that reason we used 2 fold cross validation over the training set to estimate it (a detailed description of GSFM can be found in [18]). BER is computed as follows:

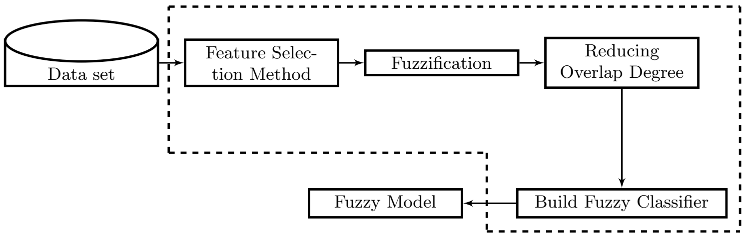

Fig. 1. Process for building a model

$$BER = \frac{1}{j}\sum_{i=1}^{j} e_i \qquad (1)$$

where j is the number of classes in the data set and e_i is the error in the i^{th} class.

Table 2. Methods considered in GSFM. The combination of feature selection method, type of fuzzy processing, and fuzzy classifier is done with a genetic algorithm.

Method	Number of Parameters	Description
Feature Selection		
ReliefF	3	Ranking of features based on the ReliefF algorithm.
χ^2	1	Ranking of features based on the χ^2 statistical test.
InfoGain	1	Features are ranked according to their information gain.
Correlation	2	A subset of features is selected according to the correlation among themselves.
Fuzzy Processing		
NLP	1	Defines the number of linguistic properties. It can take the values: 3 (Low, Medium, and High), 5 (Very Low, Low, Medium, High, and Very High), and 7 (Very Low, Low, More or Less Low, Medium, More or Less High, High, and Very High).
TMF	1	The type of fuzzy membership function. It can take the values of Trapezoid, Triangle, Gaussian, and Bell.
Fuzzy Classifier		
FDT	1	A fuzzy decision tree.
FDF	2	A fuzzy decision forest.
FKNN	2	A fuzzy version of the k nearest neighbor algorithm.
FRNN	5	A fuzzy relational neural network classifier.

4 Experiments and Results

For our experiments we used the Baby chillanto® infant cry data base property of INAOE-CONACyT, Mexico. This data set has samples of cries from deaf babies and with asphyxia, hunger, pain, and normal, as described in section 2.

We performed several experiments considering binary classification. First, we considered the binary problems identifying between asphyxia and normal[1] cries, deaf and normal, and finally hungry and pain cries. For each experiment, we used GSFM to determine the best model.

The evaluation was done using 10 fold cross validation. This technique divides the data set into 10 disjoint subsets, and in each fold a subset is left apart for testing and the remaining subsets for training. This process is repeated until all subsets have been used for testing and training. As evaluation metrics we used accuracy (ACC), true positive rate (TPR), true negative rate (TNR) and area under the ROC[2] curve (AUC) [4].

Table 3 shows the best obtained results in our experiments and the reported by Rosales-Pérez et al. [17]. Even though other works have tackled the infant cry classification problem, a direct comparison among those is not performed, because they do not apply 10 fold cross validation. Nonetheless, accuracy percentages of 99% and of 95% are reported in [9] and [19], respectively, for asphyxia vs normal cries. Table 3 also shows that our approach clearly outperforms the reported in [17]. However, we performed the Wilcoxon signed rank test [23] to determine whether the difference is statistically significant or not. This test was done across the obtained results in each fold. Table 4 shows this test and whether the difference for each data set is significant or not.

Table 3. Percentual classification results for each experiment. Results are the average of using 10 fold cross validation. Reported results by Rosales-Pérez et al. [17] are also shown. The best result is shown in bold font for each case.

ID	Data Set	Accuracy		TPR		TNR		AUC	
		GSFM	[17]	GSFM	[17]	GSFM	[17]	GSFM	[17]
1	Asphyxia vs Normal	**90.68**	88.67	85.29	**90.00**	**94.29**	87.78	**95.79**	92.85
2	Deaf vs Normal	**99.42**	97.55	**100.00**	98.75	**98.42**	95.47	**100.00**	99.75
3	Hungry vs Pain	**97.96**	96.03	**99.43**	95.59	95.26	**96.67**	**98.89**	98.35

Table 4. Statistical test on the infant cry data set. T^+ is the sum of possitive differences, T^- is the sum of negative difference and T is the minimum between T^+ and T^-. The significance level, α, is set to 0.05, for that level T should be less than 8

ID	T^+	T^-	T	significant?
1	46	6	6	YES
2	41	7	7	YES
3	40	9	9	NO

Finally, Table 5 describes the obtained models for each case. For each model, the feature selection method (FSM), type of membership function (TMF), number of linguistic properties (NLP), the learning algorithm (LA), and its associated parameters are shown.

[1] We considered the hungry cries as normal cries.

[2] Receiver Operating Characteristic.

Table 5. Selected models for infant cry data sets using GSFM

ID	FSM	TMF	NLP	LA	Parameters
1	Correlation	Trapezoid	3	FKNN	nn = 7
					sm = correlation
2	InfoGain	Bell	7	FDT	cv = 0.87
3	InfoGain	Gauss	3	FKNN	nn = 5
					sm = chord

5 Conclusion and Future Work

In this work we described the application of GSFM for the task of infant cry classification. Our approach allows to select an adequate model to differentiate between each type of cry. Among the main advantages of the adopted approach we highlight that the user does not have to perform several experiments in order to determine a good combination of methods. Our experimental results show that our approach outperforms results reported in the literature from methods that only consider one learning algorithm.

As a future work we would like to explore the use of other techniques to extract features from the audio signals, as well as to test strategies such as ensembles of different models. We will also try infant cry classification as a multiclass problem.

Acknowledgements. The firts author (register number 234672) wants to thank to CONACyT for the financial support given to his postgraduate studies.

References

1. Barajas, S., Reyes, C.: Your fuzzy relational neural network parameters optimization with a genetic algorithm. In: The 14th IEEE International Conference on Fuzzy Systems, FUZZ 2005, pp. 684–689. IEEE (2005)
2. Barajas-Montiel, S., Reyes-García, C.: Fuzzy Support Vector Machines for Automatic Infant Cry Recognition. In: Huang, D.-S., Li, K., Irwin, G. (eds.) Intelligent Computing in Signal Processing and Pattern Recognition. LNCIS, vol. 345, pp. 876–881. Springer, Heidelberg (2006)
3. Barajas-Montiel, S.E., Reyes-García, C.A., Arch-Tirado, E., Mandujano, M.: Improving Baby Caring with Automatic Infant Cry Recognition. In: Miesenberger, K., Klaus, J., Zagler, W.L., Karshmer, A.I. (eds.) ICCHP 2006. LNCS, vol. 4061, pp. 691–698. Springer, Heidelberg (2006)
4. Beck, J., Shultz, E., et al.: The use of relative operating characteristic (roc) curves in test performance evaluation. Archives of Pathology & Laboratory Medicine 110(1), 13 (1986)
5. Boersma, P., Weenink, D.: Praat v. 4.0. 8. A system for doing phonetics by computer. Institute of Phonetic Sciences of the University of Amsterdam (2002)

6. Cano Ortiz, S.D., Escobedo Beceiro, D.I., Ekkel, T.: A Radial Basis Function Network Oriented for Infant Cry Classification. In: Sanfeliu, A., Martínez Trinidad, J.F., Carrasco Ochoa, J.A. (eds.) CIARP 2004. LNCS, vol. 3287, pp. 374–380. Springer, Heidelberg (2004)
7. Engelbrecht, A.: Computational intelligence: an introduction. Wiley (2007)
8. Galaviz, O.F.R., García, C.A.R.: Infant Cry Classification to Identify Hypo Acoustics and Asphyxia Comparing an Evolutionary-Neural System with a Neural Network System. In: Gelbukh, A., de Albornoz, Á., Terashima-Marín, H. (eds.) MICAI 2005. LNCS (LNAI), vol. 3789, pp. 949–958. Springer, Heidelberg (2005)
9. Hariharan, M., Yaacob, S., Awang, S.A.: Pathological infant cry analysis using wavelet packet transform and probabilistic neural network. Expert Systems with Applications 38(12), 15377–15382 (2011), `http://www.sciencedirect.com/science/article/pii/S0957417411009201`
10. Holland, J.: Adaptation in natural and artificial systems. University of Michigan Press, Ann Arbor MI (1975)
11. Lederman, D.: Automatic classification of infants cry. Ben Gorion University. M.Sc. Thesis P, pp. 1–11 (2002)
12. Lederman, D., Cohen, A., Zmora, E., Wermke, K., Hauschildt, S., Stellzig-Eisenhauer, A.: On the use of hidden markov models in infants' cry classification. In: The 22nd Convention of Electrical and Electronics Engineers in Israel, pp. 350–352. IEEE (2002)
13. Lederman, D., Zmora, E., Hauschildt, S., Stellzig-Eisenhauer, A., Wermke, K.: Classification of cries of infants with cleft-palate using parallel hidden markov models. Medical and Biological Engineering and Computing 46, 965–975 (2008)
14. Orozco-García, J., Reyes-García, C.A.: A Study on the Recognition of Patterns of Infant Cry for the Identification of Deafness in Just Born Babies with Neural Networks. In: Sanfeliu, A., Ruiz-Shulcloper, J. (eds.) CIARP 2003. LNCS, vol. 2905, pp. 342–349. Springer, Heidelberg (2003)
15. Reyes-Galaviz, O.F., Tirado, E.A., Reyes-Garcia, C.A.: Classification of Infant Crying to Identify Pathologies in Recently Born Babies with ANFIS. In: Miesenberger, K., Klaus, J., Zagler, W.L., Burger, D. (eds.) ICCHP 2004. LNCS, vol. 3118, pp. 408–415. Springer, Heidelberg (2004)
16. Reyes-Galaviz, O.F., Verduzco, A., Arch-Tirado, E., Reyes-García, C.A.: Analysis of an Infant Cry Recognizer for the Early Identification of Pathologies. In: Chollet, G., Esposito, A., Faúndez-Zanuy, M., Marinaro, M. (eds.) Nonlinear Speech Modeling. LNCS (LNAI), vol. 3445, pp. 404–409. Springer, Heidelberg (2005)
17. Rosales-Pérez, A., Reyes-García, C.A., Gómez-Gil, P.: Genetic Fuzzy Relational Neural Network for Infant Cry Classification. In: Martínez-Trinidad, J.F., Carrasco-Ochoa, J.A., Ben-Youssef Brants, C., Hancock, E.R. (eds.) MCPR 2011. LNCS, vol. 6718, pp. 288–296. Springer, Heidelberg (2011)
18. Rosales-Pérez, A., Reyes-García, C.A., Gómez-Gil, P., Gonzalez, J.A., Altamirano, L.: Genetic Selection of Fuzzy Model for Acute Leukemia Classification. In: Batyrshin, I., Sidorov, G. (eds.) MICAI 2011, Part I. LNCS, vol. 7094, pp. 537–548. Springer, Heidelberg (2011)
19. Sahak, R., Mansor, W., Lee, Y.K., Yassin, A.I.M., Zabidi, A.: Performance of combined support vector machine and principal component analysis in recognizing infant cry with asphyxia. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6292–6295. IEEE (2010)

20. Suaste-Rivas, I., Díaz-Méndez, A., Reyes-García, C.A., Reyes-Galaviz, O.F.: Hybrid Neural Network Design and Implementation on FPGA for Infant Cry Recognition. In: Sojka, P., Kopeček, I., Pala, K. (eds.) TSD 2006. LNCS (LNAI), vol. 4188, pp. 703–709. Springer, Heidelberg (2006)
21. Suaste-Rivas, I., Reyes-Galaviz, O.F., Diaz-Mendez, A., Reyes-Garcia, C.A.: A Fuzzy Relational Neural Network for Pattern Classification. In: Sanfeliu, A., Martínez Trinidad, J.F., Carrasco Ochoa, J.A. (eds.) CIARP 2004. LNCS, vol. 3287, pp. 358–365. Springer, Heidelberg (2004)
22. Suaste-Rivas, I., Reyes-Galviz, O.F., Diaz-Mendez, A., Reyes-Garcia, C.A.: Implementation of a Linguistic Fuzzy Relational Neural Network for Detecting Pathologies by Infant Cry Recognition. In: Lemaître, C., Reyes, C.A., González, J.A. (eds.) IBERAMIA 2004. LNCS (LNAI), vol. 3315, pp. 953–962. Springer, Heidelberg (2004)
23. Wilcoxon, F.: Individual comparisons by ranking methods. Biometrics Bulletin 1(6), 80–83 (1945)