Biomedical Diagnosis of Infant Cry Signal Based on Analysis of Cepstrum by Deep Feedforward Artificial Neural Networks

Salim Lahmiri, Chakib Tadj, and Christian Gargour

The automatic analysis and detection of audio signals is an important field of research with promising applications in various biomedical engineering problems such as speech, heart murmur, and lung sound analysis and classification. In this regard, automatic classification of infant vocalizations is becoming an appealing research area for medical diagnosis in clinical milieu. Indeed, the analysis and classification of infant cry records is a conventional non-invasive technique to distinguish between healthy and unhealthy infants.

Recently, various computer-aided diagnosis (CAD) systems have been developed to detect infant pathological cry records. For instance, an automatic segmentation system for newborn cry recordings was developed in [1], based on Hidden Markov Models to detect cry expiratory and inspiratory parts from normal and pathological newborns. The proposed system achieved 83.79% accuracy. The authors in [2] used a features vector composed of the prevalence of fundamental frequency glide, resonance frequencies dysregulation, and Mel-frequency cestrum coefficients to train a probabilistic neural network. The latter achieved 88.71% accuracy in classifying healthy and unhealthy records of preterm babies and achieved 67.00% accuracy in classifying healthy and unhealthy records of full-term babies. More recently, the authors in [3] proposed an automatic system that combines short-term and long-term features from different time scales to distinguish between the cry audio signals of healthy infants from those with respiratory distress syndrome. When trained with Mel-frequency cepstral coefficients, tilt, and rhythm features, the linear support vector machine yielded to 73.80% accuracy when tested on expiration samples and 67.80% accuracy when tested on inspiration samples.

In this work, we propose a new CAD system to distinguish between healthy and unhealthy infant cry signals. The proposed CAD system is composed of four major steps. First, the original cry signal is pre-processed to remove background noise and artifacts. This step also includes signal segmentation to differentiate between expiration and segmentation episodes. Second, the resulting pre-processed cry signal is analyzed to obtain its cepstrum. Third, the obtained cepstrum coefficients are fed to a deep feedforward neural network (DFFNN) for training and classification. Fourth, the performance of the cepstrum-DFFNN system is evaluated by standard classification performance metrics.

The cepstrum is widely employed in audio signal analysis as it provides a description of the spectrum envelope and spectral richness and characterizes the harmonic and noise components of the original signal [4]. Moreover, in recent years, there has been a growing interest in deep learning in various engineering and science problems thanks to its ability to extract deep features and achieve high accuracy compared to existing machining learning techniques. However, the application of deep learning to the problem of infant cry signal classification for medical diagnosis has not been explored in the biomedical literature, except in the classification of baby cry signals under different domestic environment conditions [5]. In this work, we focus on deep feedforward neural network as it has deeper architectures compared to standard feedforward neural network, which allow the input data to be analyzed and transformed multiple times to generate the output [6]. In this regard, multiple hidden layers make the DFFNN more appropriate for comprehensive data [7], [8]. In addition, DFFNN is faster compared to most common deep learning artificial neural network such as convolutional neural network and long short-term memory network.

The proposed CAD system for infant cry signal analysis and classification is shown in Fig. 1, in which the original cry signal is denoised and segmented. Then, it is processed to obtain its ceptrum signature. The latter is employed to train a DFFNN which is used to distinguish between healthy and unhealthy infant cry signals. For comparison purposes, Naïve Bayes, support vector machine, and probabilistic neural

This research is partly supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) [RGPIN-2016-05067].

1094-6969/21/$25.00©2021IEEE

Authorized licensed use limited to: Bibliothèque ÉTS. Downloaded on April 20,2021 at 22:20:19 UTC from IEEE Xplore. Restrictions apply.

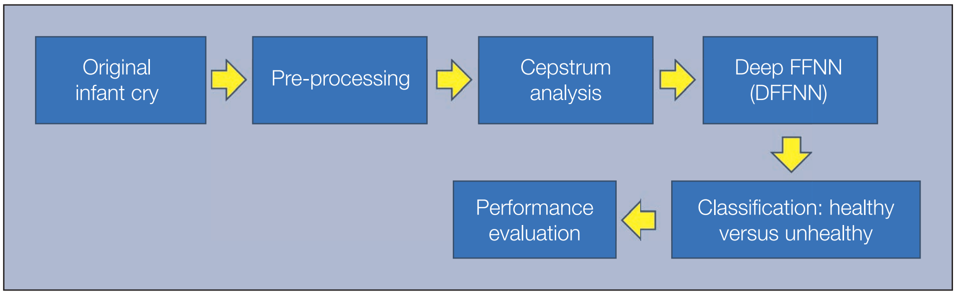

Fig. 1. Flowchart of the proposed CAD system based on cepstrum analysis and DFFNN for infant cry signal classification. The pre-processing step includes signal denoising, artifact removing, and segmentation to separate expiration and inspiration episodes. The resulting classification made by DFFNN is evaluated by standard performance measures.

network are employed as baseline classifiers. In this regard, we seek to show that the DFFNN is effective compared to the aforementioned, and investigation of various combinations of DFFNN architectures and related functions are out of scope of the current work. The proposed CAD system for infant cry analysis and classification will be applied to a large data set obtained from a Canadian hospital located in Montreal, Quebec and two other hospitals located in Lebanon.

Cepstrum Analysis

Basically, the cepstrum is the inverse Fourier transform of the logarithm of the signal spectrum. In other words, it is essentially a spectrum of the original signal spectrum. The most compelling feature of cepstrum is that any periodic pattern in the spectrum arises as a particular component of cepstrum [9]. In this regard, cepstrum analysis has been employed in various biomedical signal processing applications, including epileptic seizure detection [9], assessing Parkinson's disease severity [10], newborn cry diagnostics [3], classification of heart sounds [11], and estimation of heartbeat rate [12]. Besides, cepstrum analysis was successfully employed in mechanical fault diagnosis [13]–[15].

Technically speaking, the cepstrum $C[n]$ is computed as the inverse discrete Fourier transform (IDFT) of the log magnitude of the DFT of a signal $x[n]$, which is given as follows:

$$C\left[n\right] = IDFT\left\{\log\left|DFT\left(x\left[n\right]\right)\right|\right\} \tag{1}$$

Keep in mind that there is no formal guide on how to choose the number of coefficients used to describe the cepstrum. In this study, the number of cepstrum coefficients to estimate is set to 1000. Indeed, we make the hypothesis that such number can statistically describe harmonics in the original infant cry signal. Also, we expect that the DFFNN would take less time to converge when trained with a feature vector of size 1000 in each single hidden layer.

Deep Feedforward Neural Network

The standard feedforward neural network (FFNN) is an artificial neural network with one hidden layer used to process the inputs. Besides, the deep feedforward neural network (DFFNN) has several hidden layers. Specifically, the information in DFFNN moves from the input layer through the hidden layers to the output layer, and there is no feedback or loop in the network [16] which basically makes it deep and fast. Recall that deep learning artificial neural networks have been found to be successful in environment sound classification [17], improving safety of elderly people [18], heart rate estimation [19], environmental and biological measurement [20], and hand gesture recognition [21].

The architecture of FFNN used for infant cry classification is presented in Fig. 2. Accordingly, there are one input layer, three hidden layers, and one output layer. The number of neurons is set to 1000 in the input layer and in each hidden layer. The number of neurons in the output layer is set to one to represent the class label: either a healthy cry record or an unhealthy cry record. Each hidden neuron is used to process output information of the input layer according to the following expression:

$$h\left(X\right) = wX + b \tag{2}$$

where X is the input vector, w is the matrix of weights, and b is the bias vector. In general, the neuron of the output is nonlinearly processed by a given activation function such as sigmoid, tanh, ReLU, and ELU functions. For instance, the input $h(X)$ is

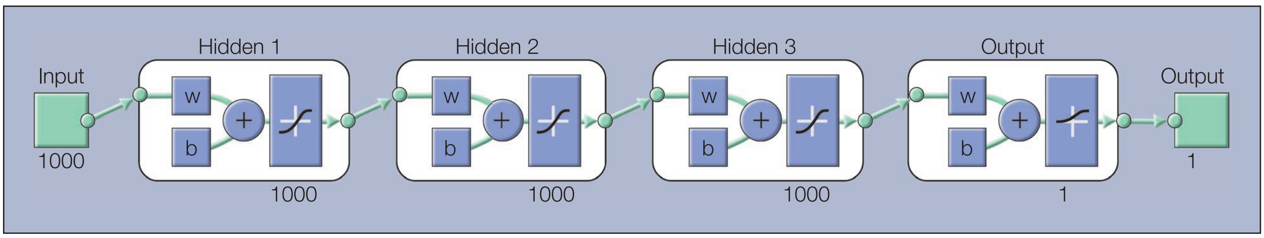

Fig. 2. DFFNN with three hidden layers. The input layer has 1000 neurons, corresponding to the number of coefficients in the cepstrum. There are 1000 neurons in each hidden layer. The network is fully connected. Specifically, every single neuron in a layer is connected to all neurons in the next layer. As a result, there is one bias plus 1000*1000 connections in each hidden layer. The output layer has only one neuron that indicates the class label. In each hidden layer and in the output layer, *W* indicates the corresponding weight matrix and the corresponding constant parameter *b*. The activation function used to process the output signal from each layer is the sigmoid.

Authorized licensed use limited to: Bibliothèque ÉTS. Downloaded on April 20,2021 at 22:20:19 UTC from IEEE Xplore. Restrictions apply.

processed by the nonlinear activation function. In our study, the sigmoid function is chosen as the activation function and it is given by:

$$Sigmoid(x) = \frac{1}{1+e^{-x}} \tag{3}$$

Finally, the output neuron is expressed as follows:

$$y = Sigmoid\left(\sum h(X)\right) \tag{4}$$

Recall that the information fed to each neuron in each hidden layer is processed by a sigmoid activation function. The resilient backpropagation algorithm is chosen as the training function since it is fast and does not require large memory during computation. Finally, the number of epochs is set to 10, and the learning rate is set to 0.01.

Recall that there is no rule of thumb on how to configure the entire architecture of the DFFNN along with its parameters. In general, they are fixed, according to the experience of the user. In our study, the number of hidden layers is set to four for deeper analysis of the cepstum coefficients compared to two and three hidden layers and to obtain acceptable computation processing time. Also, the number of epochs is set to 10 for fast processing of the input signal. Finally, a learning rate of 0.01 is a good compromise between convergence and required processing time. Indeed, our goal is to design an effective and fast DFFNN system for analysis and classification of ceptrum coefficients.

Baseline Classifiers: SVM, NB, and PNN

The SVM classifier [22] is a supervised learning algorithm based on statistical learning theory used to determine a hyper plane. The latter optimally splits two classes by learning a train data set. For instance, let $\{x_i, y_i\}_{i=1}^{n}$ where x is the input vector, and y is the class label. The classification decision function is expressed as follows:

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x_j) + b\right) \tag{5}$$

where α_i is the Lagrange multipliers, $K(x_i,x_j)$ is a linear kernel function, and b is a constant parameter.

The Naïve Bayes classifier [23] is based on estimation of probabilities to assign the membership of an input vector x to a particular class y. The Bayes' theorem can be used to express the conditional probability of class label y given input vector x as follows:

$$P(y \mid x) = \frac{P(x \mid y) P(y)}{P(x)} \tag{6}$$

where $P(y \mid x)$ is the posterior probability of the compound class, $P(x \mid y)$ is the conditional probability that a compound has certain features given its class y, $P(y)$ is the prior probability estimated from the training set, and $P(x)$ is the marginal probability of observing the given features in the dataset.

The probabilistic neural network (PNN) [24] is composed of three layers: the input layer, pattern layer, summation layer, and output layer. The input layer has 1000 neurons, corresponding to the number of cepstrum coefficients. The second layer consists of 1000 neurons where each one is represented by a Gaussian transfer function. The pattern layer has two neurons where each one is used to represent a specific class. The pattern layer is used to perform an average operation of the outputs from the pattern layer for each class. Finally, the output layer has one neuron used to compute the maximum sum as follows:

$$y = argmax_i \left\{P_i(y)\right\} \tag{7}$$

where,

$$P_i(y) = \frac{1}{N_i} \sum_{j=1}^{N_i} Q_{ij}(y) \tag{8}$$

and,

$$Q_{ij}(y) = \frac{1}{(2\pi)^{0.5}\,\delta} exp\left(\frac{-\left(y - x_{ij}\right)^T \left(y - x_{ij}\right)}{2\delta^2}\right) \tag{9}$$

where $P_i(y)$ represents the probability that a test sample x belongs to class y, Q_{ij} represents the standard probability density function (PDF), δ is a smoothing parameter set to unity, and N is the number of neurons in the input layer.

Experimental Protocol and Performance Measures

The performance of each classifier is measured by computing the accuracy, sensitivity, and specificity. The accuracy is the ratio of the correct predictions to the total number of predictions, sensitivity measures the proportion of positive predictions that are correctly identified over all positive cases, and specificity measures the proportion of negative predictions that are correctly identified. Hence, accuracy, sensitivity, and specificity are expressed as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{10}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{11}$$

$$Specificity = \frac{TN}{TN + FP} \tag{12}$$

where TP, TN, FN and FP indicate the number of true positives, true negatives, false negatives and false positives, respectively.

To evaluate the performance of each classifier while avoiding overfitting, 10-fold cross-validation protocol is adopted in our study. For instance, under 10-fold cross-validation protocol, the data set is divided into 10 subsets. Each time, one different subset is used as the test set, and the remaining nine subsets are put together to form a training set. Then the average performance across all 10 trials is calculated. In the current work, the average and standard deviation of accuracy, sensitivity, and specificity are calculated across the 10 folds. Also,

Authorized licensed use limited to: Bibliothèque ÉTS. Downloaded on April 20,2021 at 22:20:19 UTC from IEEE Xplore. Restrictions apply.

we consider an additional experimental protocol where the data is randomly separated into 50% for learning and 50% for testing.

Data and Pre-processing

The database is composed of two sets: expiration (EXP) set and inspiration (INS) set. The EXP set has 2638 cry signals and INS set has 1860 cry signals. Specifically, there are 1319 healthy signals and 1319 unhealthy signals in the EXP set. Moreover, there are 930 healthy signals and 930 unhealthy signals in the INS set. To record a cry signal, a two-channel sound recorder with a sampling frequency of 44.1 kHz and a resolution of 16 bits was placed at 10 cm to 30 cm from the infant. The time duration of each recorded signal is within 2 to 3 minutes. Each original recorded cry signal has been pre-processed to remove background noise and artifacts. It is also segmented to keep only respiration and expiration episodes. The segmentation task is manually performed by using the Wave Surfer tool.

All infant cry signals have been recorded in the neonatology departments of the following hospitals: Sainte-Justine hospital (Montreal, Canada), Al-Sahel hospital (Beirut, Lebanon), and Al-Raee hospital (Saida, Lebanon). The infants who entered the study are preterm and full term, and their respective ages range from one to 53 days. The sample includes both healthy and unhealthy babies and both males and females. The group of unhealthy babies suffers from various pathologies such as diseases affecting the central nervous system and respiratory system. Other pathologies include blood disorders, chromosomal abnormalities, and congenital cardiac anomalies. For illustration purpose, Fig. 3 displays examples of healthy and unhealthy signals.

Examples of cepstrums representing healthy and unhealthy cry signals are shown in Fig. 4. Recall that the number of coefficients to be calculated was set to 1000 to allow for a general representation of the original cry signal, on one hand, and for fast learning and classification by each single classifier, on the other hand. According to Fig. 4, INS cepstrums show

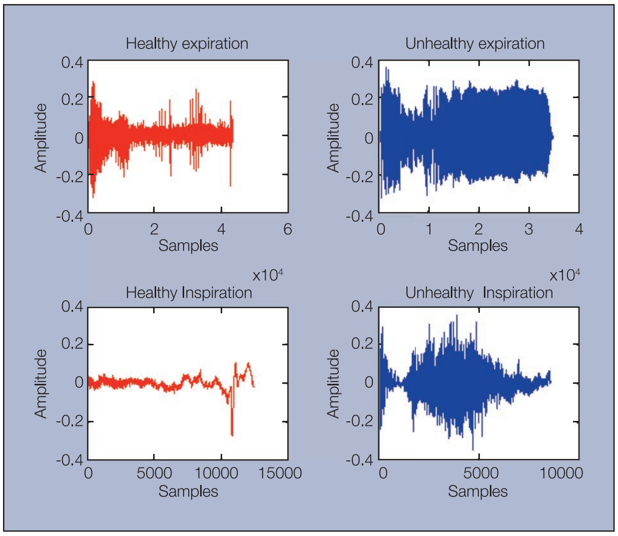

Fig. 3. Examples of healthy and unhealthy cry signals.

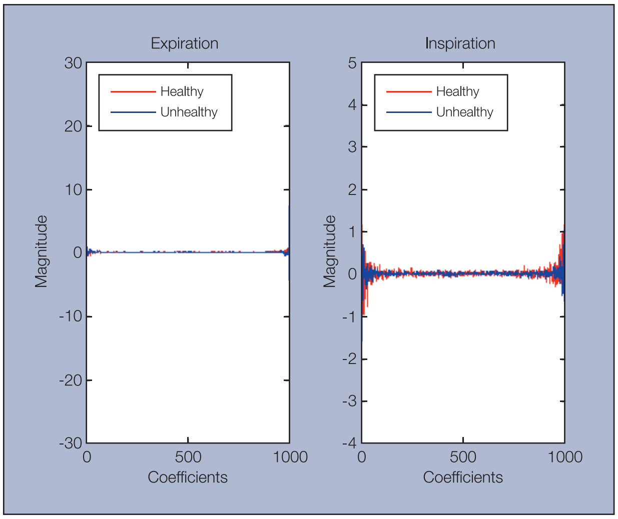

Fig. 4. Examples of cepstrums from healthy and unhealthy infant cry signals. More variability is observed in cepstrums from healthy and unhealthy signals associated with inspiration records. In addition, visual inspection of cepstrum coefficients from inspiration set indicates that those related to healthy infant cry signals exhibit higher magnitude compared to those associated with unhealthy cry signals.

more variability compared to those from EXP sets. All experiments are performed by a PC with 1.80 GHz processor and 8 GB installed RAM in Matlab2020b© cloud environment.

Experimental Results

Fig. 5a and Fig. 5b compare the classification results from deep feedforward (DFFNN), linear support vector machine (SVM), Naïve Bayes (NB), and probabilistic neural network (PNN) following 10-fold cross-validation protocol when applied to EXP and INS sets, respectively. For both EXP and INS sets, the DFFNN outperforms the linear SVM, NB, and PNN in terms of accuracy, sensitivity, and specificity.

Specifically, in the problem of classifying EXP cry signals, the DFFNN, SVM, NB, and PNN achieved an accuracy of 99.92%±0.00, 61.15%±0.04, 58.11%±0.01, and 56.71%±0.01, respectively. In addition, the obtained sensitivity is 99.85%±0.00, 61.03%±0.04, 56.31%±0.01, and 57.70%±0.03, respectively, for DFFNN, SVM, NB, and PNN. In terms of specificity, the DFFNN, SVM, NB, and PNN obtained 100%, 61.27%±0.05, and 59.93%±0.02, and 55.72%±0.02, respectively.

Moreover, in the problem of classifying INS cry signals, the DFFNN achieved perfect accuracy, sensitivity, and specificity. The linear SVM, NB, and PNN, respectively, achieved an accuracy of 59.57%±0.01, 55.46%±0.02, and 52.63%±0.05, a sensitivity of 58.82%±0.04, 55.84%±0.01, and 47.10%±0.07, and a specificity of 60.32%±0.04, 55.08%±0.02, and 58.17%±0.06.

Finally, Table 1 provides the performance measures of each predictive model when the data set is randomly split into 50% for training and 50% for testing. As shown, the DFFNN outperforms all classifiers on both EXP set INS sets.

In summary, the DFFNN performs best in distinguishing between healthy and unhealthy infant cry signals. It is followed by the SVM, and the NB classifier performs the worst. In terms of processing time for training and classifying EXP and

Authorized licensed use limited to: Bibliothèque ÉTS. Downloaded on April 20,2021 at 22:20:19 UTC from IEEE Xplore. Restrictions apply.

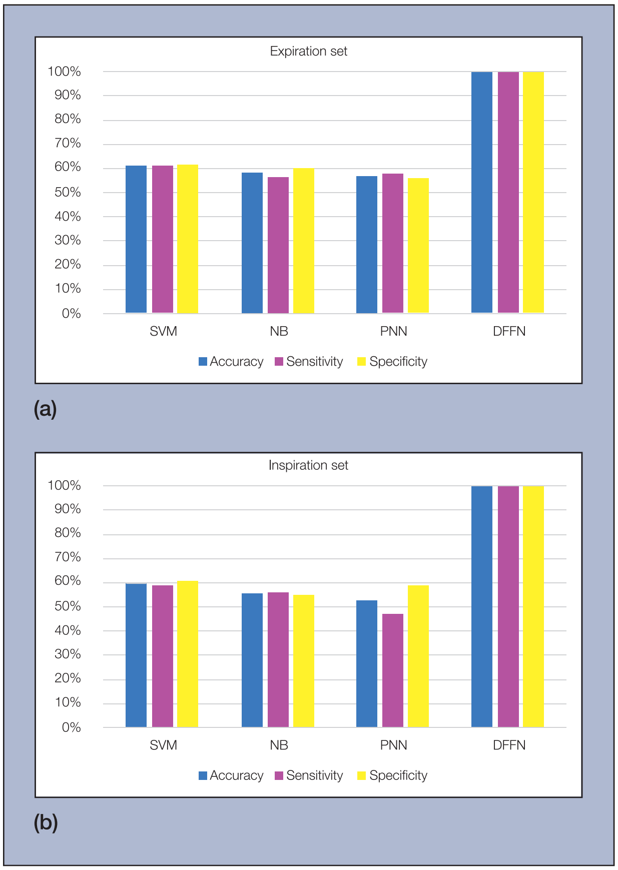

Fig. 5. Bar plots of obtained performance measures from 10-fold cross-validation protocol. The DFFNN outperforms the linear SVM, NB, and PNN classifiers in terms of accuracy, sensitivity, and specificity. Bar plots of obtained performance measures from expiration set under 10-fold cross-validation protocol. (a) Expiration set. (b) Inspiration set.

INS sets under 10-fold cross-validation protocol, the NB is the fast one, followed by PNN, DFFNN and SVM, respectively, as shown in Table 2.

The DFFNN outperforms all classifiers and requires less than one minute to converge. Its superiority can be explained by the ability to capture complexities inherent in the data structure. Specifically, as they have high level of abstraction, the DFFNN can fully account for the complex relationships between cepstrums and their corresponding classes. Moreover, the slight improvement in performance measures observed for the INS set that yield to perfect accuracy by DFFNN could be attributed to its capability to learn and model large variability in cepstrums, which is clearly observed in Fig. 4.

Table 2 – Processing time (in seconds) of each classifier under 10-fold cross-validation

	Expiration (EXP) set	Inspiration (INS) set
SVM	653.14	73.57
NB	3.18	1.46
PNN	19.57	14.30
DFFNN	29.30	31.63

In addition, it is worth mentioning that the NB classifier is fast since it requires only a small training data set to estimate the probabilities. However, it performed the worst, which is most likely linked to violation of the assumption of independent predictors, for instance, cepstrum coefficients. Furthermore, it is worth mentioning that the performance of the linear SVM is moderate, possibly because its key parameters have not been optimized, and it is very slow compared to NB and DFFNN due to the complexity of the optimization process that underlies the SVM and the complexity of the data under study. The accuracy of the PNN is the lowest, and this could be explained by the fact that this kind of artificial neural network employs a Gaussian transfer function to process the inputs. In this regard, its performance depends on the distribution of the inputs and on the value of the smoothing parameter used to determine the length of the probability density function (PDF) of the transfer function.

Finally, it is worth mentioning that the DFFNN outperformed most recent studies in distinguishing between healthy and unhealthy infant cry signals in terms of accuracy, including Hidden Markov Models trained with segmented cry signals (83.79%) [1], probabilistic neural network trained with prevalence of fundamental frequency glide, resonance frequencies dysregulation, and Mel-frequency cestrum coefficients (67.00% to 88.71%) [2], and linear SVM trained with Mel-frequency cepstral coefficients, tilt, and rhythm features, (67.80%) [3]. Therefore, the proposed CAD system for infant cry signal classification based on DFFNN trained with cepstrum coefficients appears to be effective and promising.

Conclusion

Infant cry signal analysis is a non-invasive acoustic evaluation that represents an important tool for physicians for pathology diagnosis. The purpose of the current work was to design a CAD system to distinguish between healthy and unhealthy infant cry signals. In this regard, we proposed to calculate the

Table 1 – Experimental results from random split of the data into 50% learning and 50% testing

	Accuracy	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity
	Expiration set			Inspiration set		
SVM	56.63%	59.79%	53.48%	54.62%	51.61%	57.63%
NB	57.32%	55.91%	58.73%	57.10%	55.48%	58.71%
PNN	53.83%	52.88%	54.78%	52.80%	44.73%	60.86%
DFFNN	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Authorized licensed use limited to: Bibliothèque ÉTS. Downloaded on April 20,2021 at 22:20:19 UTC from IEEE Xplore. Restrictions apply.

cepstrum of infant cry signal to capture its periodic patterns. Then, the obtained cepstrum coefficients were fed to deep feedforward neural network for training and classification under ten-fold cross-validation protocol.

Thanks to its ability to learn and model deep complex structures in the data, the deep feedforward neural network achieved very close to perfect accuracy when applied to expiration infant cry signals and yielded to perfect accuracy when applied to inspiration infant cry signals. In addition, it outperformed the linear SVM and the Naïve Bayes systems when tested both on the expiration and inspiration sets. Furthermore, the proposed approach outperformed very recent works found in the literature. In short, the proposed system for infant cry signal analysis and classification was found to be effective and fast.

In future work, some interesting issues will be considered. First, regarding DFFNN, various topologies and transfer functions will be evaluated. As this is out of scope of the current work, such investigation is expected to be comprehensive regarding the possible architectures, parameters, and transfer functions. Second, other categories of deep learning neural networks will be examined and compared to DFFNN, such as convolutional neural network and long short-term memory networks. Third, we will examine the performance of optimized classifiers when trained with two types of features, namely deep features and nonlinear statistical features. Such investigation will shed light on the effectiveness of other existing deep learning neural networks and on the discrimination power of deep features compared to features used to describe nonlinear dynamics in infant cry signals.

References

[1] L. Abou-Abbas, H. Fersaie Alaie, and C. Tadj, "Automatic detection of the expiratory and inspiratory phases in newborn cry signals," *Biomedical Signal Processing and Control*, vol. 19, pp. 35-43, 2015.

[2] Y. Kheddache and C. Tadj, "Identification of diseases in newborns using advanced acoustic features of cry signals," *Biomedical Signal Processing and Control*, vol. 50, pp. 35-44, 2019.

[3] F. Salehian Matikolaie and C. Tadj, "On the use of long-term features in a newborn cry diagnostic system," *Biomedical Signal Processing and Control*, vol. 59, 101889, 2020.

[4] C. M. Travieso, J. B. Alonso, M. del Pozo, J. R. Ticay, and G. Castellanos-Dominguez, "Building a cepstrum-HMM kernel for apnea identification," *Neurocomputing*, vol. 132, pp. 159-165, 2014.

[5] R. Cohen, D. Ruinskiy, J. Zickfeld, H. IJzerman, and Y. Lavner, "Baby cry detection: deep learning and classical approaches," in *Development and Analysis of Deep Learning Architectures, Studies in Computational Intelligence*, W. Pedrycz and S. M. Chen, Eds. Cham, Switzerland: Springer, vol. 867, 2020.

[6] X. J. Luo, L. O. Oyedele, A. O. Ajayi, O. O. Akinade, J. M. Davila Delgado, H. A. Owolabi, and A. Ahmed, "Genetic algorithm-determined deep feedforward neural network architecture for predicting electricity consumption in real buildings," *Energy and AI*, vol. 2, 100015, 2020.

[7] J. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Networks*, vol. 61, pp. 85-117, 2015.

[8] Y. LeCun, Y. Bengio, and G Hinton, "Deep learning," *Nature*, vol. 321, pp. 436-444, 2015.

[9] E. Yavuz, M. C. Kasapbaşi, C. Eyüpoğlu, and R. Yazıcı, "An epileptic seizure detection system based on cepstral analysis and generalized regression neural network," *Biocybernetics and Biomedical Engineering*, vol. 38, pp. 201-216, 2018.

[10] T. Khan, L. E. Lundgren, D. G. Anderson *et al*., "Assessing Parkinson's disease severity using speech analysis in non-native speakers," *Computer Speech Language*, vol. 61, 101047, 2020.

[11] Z. Abduh, E. A. Nehary, M. Abdel Wahed, and Y. M. Kadah, "Classification of heart sounds using fractional Fourier transform based mel-frequency spectral coefficients and traditional classifiers," *Biomedical Signal Processing and Control*, vol. 57, 101788, 2020.

[12] H. Ankıshan, "Estimation of heartbeat rate from speech recording with hybrid feature vector (HFV)," *Biomedical Signal Processing and Control*, vol. 49, pp. 483-492, 2019.

[13] P. V. Kane and A. B. Andhare, "Critical evaluation and comparison of psychoacoustics, acoustics and vibration features for gear fault correlation and classification," *Measurement*, vol. 154, 107495, 2020.

[14] Z. Fan and H. Li, "A hybrid approach for fault diagnosis of planetary bearings using an internal vibration sensor," *Measurement*, vol. 64, pp. 71-80, 2015.

[15] C. Guo, "Study on the recognition of aero-engine blade-casing rubbing fault based on the casing vibration acceleration," *Measurement*, vol. 65, pp. 71-80, 2015.

[16] N. Ketkar, "Feed forward neural networks," in *Deep Learning with Python*. Berkeley, CA, USA: Apress, pp. 17-33, 2017.

[17] X. Fan, T. Sun, W. Chen, and Q. Fan, "Deep neural network based environment sound classification and its implementation on hearing aid app," *Measurement*, vol. 159, 107790, 2020.

[18] F. Harrou, N. Zerrouki, Y. Sun, and A. Houacine, "Vision-based fall detection system for improving safety of elderly people," *IEEE Instrum. Meas. Mag*., pp. 49-55, 2017.

[19] E. Moya-Albor, J. Brieva, H. Ponce, and L. Martínez-Villaseñor, "A non-contact heart rate estimation method using video magnification and neural networks," *IEEE Instrum. Meas. Mag*., pp. 56-62, 2020.

[20] B. Dziadak, A. Jósko, and Ł. Makowski, "On nonlinear methods in environmental and biological measurements," *IEEE Instrum. Meas. Mag*., pp. 87-93, 2020.

[21] C. Nuzzi, S. Pasinetti, M. Lancini, F. Docchio, and G. Sansoni, "Deep learning-based hand gesture recognition for collaborative robots," *IEEE Instrum. Meas. Mag*., pp. 44-51, 2019.

[22] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer Verlag Inc, 1995.

[23] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd Ed. New York, NY, USA: Springer, 2008.

[24] D. F. Specht, "Probabilistic neural networks," *Neural Networks*, vol. 3, pp. 109-118, 1990.

Salim Lahmiri (salim.lahmiri@concordia.ca) is Assistant Professor in the Supply Chain and Business Technology Management Department at Concordia University, Montreal, Canada. He holds a Ph.D degree in computer science

continued on page 44

Authorized licensed use limited to: Bibliothèque ÉTS. Downloaded on April 20,2021 at 22:20:19 UTC from IEEE Xplore. Restrictions apply.