

LB/DON/62/2020

DCM 04/57 ✓

# AN INTELLIGENT HARDWARE SYSTEM FOR REAL-TIME INFANT CRY DETECTION AND CLASSIFICATION

U.P. Pradeep Dushyantha Pathirana

179411X



LIBRARY  
UNIVERSITY OF MORATUWA, SRI LANKA  
MORATUWA

Dissertation submitted in partial fulfillment of the requirement of the  
degree of Master of Science in Artificial Intelligence

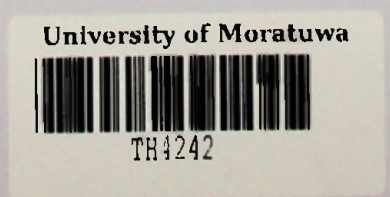
Department of Computational Mathematics

University of Moratuwa

Sri Lanka

January 2020

519.6 "2020  
004.8 (043)



TH 4242 + CD-ROM  
is not working

TH 4242

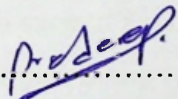
## Declaration

I declare that this dissertation does not incorporate, without any acknowledgment, any material previously submitted for a degree or a diploma in any university and to the best of my knowledge and belief, it does not contain any material previously published or written by any other person or myself except where due reference is made in the text. I also hereby give consent for my dissertation if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be available to outside organizations.

Name of Student

Signature of Student

U.P. Pradeep Dushyantha Pathirana

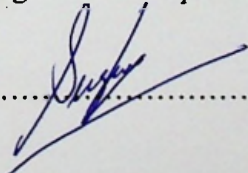
.....  


The above candidate has carried out research for the Master's dissertation under my supervision.

Supervisor

Signature of Supervisor

Dr. Sagara Sumathipala

.....  


## **Acknowledgment**

I would like to extend my heartfelt gratitude to my project supervisor Dr. Sagara Sumathipala for guiding me through the research. He extended his fullest corporation whenever I sought for advice and guidance to make this project a success. On the same note, I would also like to thank him for his guidance and support as the course coordinator during the entire course period. Similarly, I would like to extend my sincere gratitude to the entire academic staff of the Department of Computational Mathematics for their guidance throughout the course period. Further, I must thank the non-academic staff for their continuous support during the entire course in providing the learning environment to the best of their ability.

I would like to thank the entire management and research team of Synergen Technology Labs (Pvt) Ltd, for their permission to extend this project as my research project and providing me the required resources for its successful completion.

# Abstract

Cry, the universal communication language of the infants encodes vital information about the physiological and psychological health of the infant. Experienced caregivers can understand the cause of cry based on the pitch, tone, intensity, and duration. Similarly, pediatricians can diagnose hearing impairments, brain damages, and asphyxia by analyzing the cry signals, providing a non-invasive mechanism for early diagnosis in the first few months. Hence, automated cry classification has gained great importance in the fields of medicine and baby-care. With the emergence of the concept of the Internet of Things coupled with Artificial Intelligence, baby monitors have recently gained huge popularity due to features like sleep analysis, cry detection, and motion analysis through multiple sensors. Since cry classification involves audio processing in real-time, most of the solutions have either complex and costly designs or distributed computing, which leads to privacy concerns of the users. This research presents a low-cost intelligent hardware system for real-time infant cry detection and classification. The proposed solution presents the selection of the hardware to suit the requirements of audio processing while adhering to financial constraints and the firmware design, which includes voice activity detection, cry detection, and classification. This proposes the use of the multi-agent system as a resource management concept while proving that AI concepts can also be extended to resource-limited hardware platforms as the novelty. Firmware and algorithm are designed to maintain the accuracy figures above 90% while processing the audio signal at a higher rate than its production to maintain stability. A voice activity detector was designed to filter human voice through temporal features while cry detection and classification were respectively based on Artificial Neural Network and K-Nearest Neighbor algorithm trained with a spectral-domain feature vector called Mel Frequency Cepstral Coefficients (MFCC). Evaluations under diverse conditions showed accuracy figures of 96.76% and 77.45% in cry detection and classification, respectively.

**Keywords:** Cry Detection, Cry Classification, Voice Activity Detection

# Contents

1	INTRODUCTION .....	1
1.1	Prolegomena .....	1
1.2	Research Problem.....	2
1.3	Aims and Objectives .....	2
1.4	Background and Motivation.....	3
1.5	Problem in Brief .....	4
1.6	The Novelty in The Proposed Solution .....	6
1.7	Structure of the Thesis.....	6
1.8	Summary .....	7
2	CURRENT TRENDS AND ISSUES IN INFANT CRY DETECTION .....	8
2.1	Introduction .....	8
2.2	Baby Monitors.....	8
2.3	Cry Detection and Classification in Baby Monitors .....	9
2.4	Recent Works in Cry Detection and Classification.....	9
2.5	Baby Monitors with Cry Detection .....	14
2.6	Summary .....	15
3	INCORPORATED TECHNOLOGIES .....	16
3.1	Introduction .....	16
3.2	Multi-Agent Technology .....	16
3.3	Artificial Neural Network .....	19
3.4	K-Nearest Neighbour .....	21
3.5	Summary .....	22
4	HYPOTHESIS .....	23
4.1	Introduction .....	23
4.2	Hypothesis.....	23
4.3	Input.....	23
4.4	Output.....	24

4.5	Process.....	24
4.6	Users.....	24
4.7	Summary .....	25
5	HARDWARE AND FIRMWARE DESIGN.....	26
5.1	Introduction .....	26
5.2	Hardware Design.....	26
5.3	Firmware Design .....	28
5.4	Audio capturing and Noise filtering.....	30
5.5	Temporal Feature Extraction and Voice Activity Detection.....	31
5.6	Spectral Feature Extraction and Cry detection.....	31
5.7	Cry Classifier.....	32
5.8	Summary .....	32
6	IMPLEMENTATION OF CRY CLASSIFIER .....	34
6.1	Introduction .....	34
6.2	Hardware Selection .....	34
6.3	Firmware Design.....	36
6.4	Summary .....	63
7	EVALUATION OF THE SYSTEM.....	65
7.1	Introduction .....	65
7.2	Test Scenarios.....	65
7.3	Test Setup.....	65
7.4	Definitions of the Evaluation Parameters.....	66
7.5	Performance of the Cry Detector.....	68
7.6	Performance of the Cry Classifier .....	69
7.7	Summary .....	73
8	CONCLUSION AND FUTURE WORK .....	74
8.1	Introduction .....	74
8.2	Conclusion.....	74
8.3	Limitation and Related Future Works .....	74

REFERENCES.....	76
APPENDIX A: Source Codes.....	79
The Preprocessor Agent.....	79
Selection of The Best KNN Classifier for Cry Detection.....	81
Selection of The Best ANN Classifier for Cry Detection.....	82
Cry Classifier.....	83

## List of Figures

Figure 4.1 Flow diagram of the process.....	24
Figure 5.1 Hardware architecture.....	27
Figure 5.2 Firmware design .....	30
Figure 6.1 OSD3358 microprocessor.....	35
Figure 6.2 Hardware prototype .....	35
Figure 6.3 Pulse density modulated sinusoid.....	38
Figure 6.4 Frequency response of the compensation filter .....	40
Figure 6.5 Raw cry signal vs. filtered cry signal.....	40
Figure 6.6 Song filtered by compensation filter.....	41
Figure 6.7 Comparison of Raw signal vs. Down-sampled signal .....	42
Figure 6.8 Bandwidth of raw signal .....	42
Figure 6.9 Bandwidth of the down-sampled signal .....	43
Figure 6.10 Short Time Energy of cry signal.....	45
Figure 6.11 Short-Time Zero-Crossings of a cry signal .....	46
Figure 6.12 Short-Time Zero-Crossing of a noise signal.....	47
Figure 6.13 Block diagram of the Voice Activity detector.....	48
Figure 6.14 Input and output of the Voice Activity Detector .....	49
Figure 6.15 MFCC vector of the cry signal .....	49
Figure 6.16 MFCC vector of the noise signal.....	50



## List of Tables

Table 6.1 Accuracy of KNN cry detector with uniform weights .....	53
Table 6.2 Accuracy of KNN cry detector weighted with distance .....	53
Table 6.3 Accuracy of ANN cry detector with sigmoid and Tanh functions .....	55
Table 6.4 Accuracy of ANN cry detector with ReLU and Identity functions .....	55
Table 6.5 Accuracy of KNN cry classifier with uniform weights .....	58
Table 6.6 Accuracy of KNN cry classifier weighted by distance .....	59
Table 6.7 Accuracy of ANN cry classifier with sigmoid and Tanh functions.....	60
Table 6.8 Accuracy of ANN cry classifier with ReLu and Identity functions.....	60
Table 7.1 Output of the cry detector under various test scenarios .....	68
Table 7.2 Accuracy of the cry detector under various test scenarios.....	68
Table 7.3 Confusion matrix for cry frame classification .....	70
Table 7.4 Performance evaluation of cry frame classification.....	70
Table 7.5 Confusion matrix for cry-event classification.....	71
Table 7.6 Performance evaluation of cry-event classification .....	71

# List of Abbreviations

ANFIS	Adaptive Neuro-Fuzzy Inference System
AI	Artificial Intelligence
ANN	Artificial Neural Network
BFCC	Bark Frequency Cepstral Coefficients
CIC	Cascaded Integrator Comb
CNN	Convolution Neural Network
CDHMM	Continuous Density Hidden Markov Model
FFT	Fast Fourier Transform
FN	False Negative
FP	False Positive
HMM	Hidden Markov Model
IoT	Internet of Things
MLP	Multi-Layer Perceptron
KNN	K-Nearest Neighbor
LPC	Linear Predictive Coding
LPCC	Linear Predictive Cepstral Coefficients
MFCC	Mel Frequency Cepstral Coefficients
PDM	Pulse Density Modulation
PHMM	Probabilistic Hidden Markov Model
PNN	Probabilistic Neural Network
PPG	Photoplethysmogram
PRU	Programmable Realtime Unit
RBNN	Radial Basis Neural Network
RDS	Respiratory Distress Syndrome
RNN	Recurrent Neural Network
SCP	Secure Copy Protocol
SIDS	Sudden Infant Death Syndrome
SoC	System on Chip
SSH	Secure Shell
STFT	Short-Time Fourier Transform
SVM	Support Vector Machine

<b>TDNN</b>	<b>Time Delay Neural Network</b>
<b>TN</b>	<b>True Negative</b>
<b>TP</b>	<b>True Positive</b>
<b>WT</b>	<b>Wavelet Transform</b>