

ベビーベッド型デバイスを用いた乳児の啼泣認識手法の検討

角谷 星哉*¹ 藤田 和之*² 尾上 孝雄*¹ 伊藤 雄一*³

A Study on Detecting Infants' Crying Using a Baby Bed Device

Seiya Sumiya*¹, Kazuyuki Fujita*², Takao Onoye*¹ and Yuichi Itoh*³

Abstract – Recently, in Japan, people who do not have children are increasing because they feel anxious about raising their child. Hence many researchers try to eliminate one of parents' major anxieties, which is caused by unclear reasons why their baby cries. Existing research almost uses the sound of cries, but it might be highly affected by external noises. In this paper, we propose the system to obtain the data of a baby's states using a baby bed device. By utilizing data other than the sound, such as the baby's weight and the center of gravity on the bed, we try to improve accuracy of cry detection and identification the reasons of cries. Results of our early test showed that the accuracy of cry detection can be improved by combining the weight and the center of gravity with the sound, whose f-value was 0.58.

Keywords : baby cry detection, domestic environment, sound, weight, center of gravity, fast Fourier transform (FFT), support vector machine (SVM)

1. はじめに

近年、日本では少子化が深刻な問題となっている。その背景には、育児に対する不安から、子どもを持つとしない人が増えているという現状がある。その状況を改善するべく、親の育児における負担を少しでも減らし、より良い育児環境を構築する試みが広がっている。

これまでの研究により、多くの母親が育児に対して不安を抱えていることが示されている^[1]。その主な要因は、乳児への接し方に関することであり、その一つとして「なぜ泣いているのか分からない」ことが挙げられる。特に第一子を産んだばかりの母親にとって、乳児が泣いている時（啼泣時）に何をすべきか分からない、自分の行動が正しいのか分からない、といった不安は一般的であり、かつ重大な問題である。こういった不安からストレスを感じ、子供を虐待してしまうことや、子育てをしなくなってしまう可能性は十分にある。このような現状から、少子化の対策として、育児の不安を解消、あるいは削減することは重要なことであると言える。

育児不安を解消するためには、配偶者や祖父母、医者といった、育児について相談ができる相手が重要である。先行研究にて、上記のような相談相手や、その

相手の多さが育児不安の低さと大きく関係していることが示されている^[2]。乳児に対する接し方の不安を抱える親に対し、その時何をすべきか、今やっていることは正しいかなどの疑問を解消してくれる相手の存在は、育児に対する不安を削減することに直結することが分かる。

しかしながら、乳児の啼泣は発生する時間を正確に把握するのが現状困難なために、その時いつも相談できるような相手がいるとは限らない。そして、こうした状況自体が育児における不安となり得る。そのため、日常的に乳児のセンシングを行い、乳児が啼泣を始めた時にそれを認識し、その啼泣の原因を特定した上で、育児者がとるべき適切な処置を伝えるようなシステムが実現できれば、あたかも専門家が常に乳児を見守ってくれているような環境を提供でき、多くの育児者の持つ育児に対する不安を削減することができ、最終的に少子化改善の一助となると考えられる。

本稿では、まずこのシステムを実現するためのハードウェアの設計、及び製作について述べる。その後、乳児の啼泣認識において必要なデータの検討として、啼泣音声と乳児の重心・重量変化について評価する。

2. 関連研究

乳児の啼泣認識においては、啼泣の音声を用いた研究が存在する。これには例えば、メル周波数ケプトラム係数などの特徴を利用した機械学習アルゴリズムを用いた手法^[3]や、ディープスペクトラムと音響的な特徴を利用した手法^[4]、そして、連続基本周波数や高調波比等の特徴を利用した手法^[5]など、家庭環境や保育園の様な施設環境において、乳児の啼泣を音声

*1: 大阪大学 大学院情報科学研究科

*2: 東北大学 電気通信研究所

*3: 青山学院大学 理工学部

*1: Graduate School of Information Science and Technology, Osaka University

*2: Research Institute of Electrical Communication, Tohoku University

*3: College of Science and Engineering, Aoyama Gakuin University

から認識し、育児の支援に役立てようとする研究が挙げられる。また、乳児の啼泣原因の推定においては、啼泣音声の振幅スペクトルの分布の違いから、原因を「空腹」と「眠気」に分類した研究や^[6]、離散コサイン変換行列とウェーブレット変換行列を基に、音声から雑音を除去した上で、原因を「眠気」等の感情状態と「怒り・悲しみ・空腹」に分類した研究^[7]、音声の音響的特徴から、啼泣原因を「痛み」「ぐずり」「空腹」のいずれかに分類した研究^[8]、そしてメル周波数ケプストラム係数等を用いて、Dunstan Baby Languageに基づき啼泣を五種類に分類した研究^[9]など、多様なアプローチで進められている。

これらの啼泣認識と原因識別の研究は、いずれも音声を用いた手法で進められている。しかしながら、音声を利用する上で、生活音等の影響は無視できるものではなく、それを考慮した解析方法でなければ、啼泣認識や原因識別の精度が低下する可能性がある。加えて、音声を録音するための機器の設置や作動に手間がかかる場合もあり、それ自体がストレスとなる可能性も否めない。また、体重や体温の変化の様な、音声のみでは知りえない情報から、より正確な支援の提供や、啼泣時以外での見守りが実現する可能性もある。

本研究では、音声以外の乳児に関する情報を取得・解析し、音声と組み合わせる、或いは音声の代わりに用いることで、音声単体を用いるよりも雑音に強く、啼泣認識や原因推定の精度を向上させるようなシステムを提案する。具体的には、ベビーベッドに複数のセンサを取り付け、ベビーベッド上の乳児の重心・重量や深度画像等の情報を取得・解析し、乳児の啼泣の認識と原因識別を行う。このシステムは、日常的に乳児の様々な情報を取得でき、使用に関する抵抗感を最小限に抑えるよう考慮する。このシステムにより、乳児を常に見守り、啼泣時にはその原因を直ちに推定し、乳児の現在の状態とそれに応じた適切な処置を親に提供することで、育児における不安を削減する環境作りが可能となると考えられる。

3. 啼泣情報取得のためのベビーベッド型デバイス

乳児の啼泣認識、及び原因推定のためのセンシングの手法として、本研究ではベビーベッドを用いた乳児のセンシング用デバイスを提案する。乳児は、その多くの時間をベビーベッド上で生活する。そのため、そこにセンサを設置することにより、重心・重量の様な乳児のすぐそばでしか取得できないようなデータに対しても、継続的かつ自然に取得が可能となる。また、ベビーベッド自体がセンサ機器となっているため、それ以上の装置の設置や乳児への装着の手間は不要である。このデバイスは、乳児や育児者に対する特殊な機

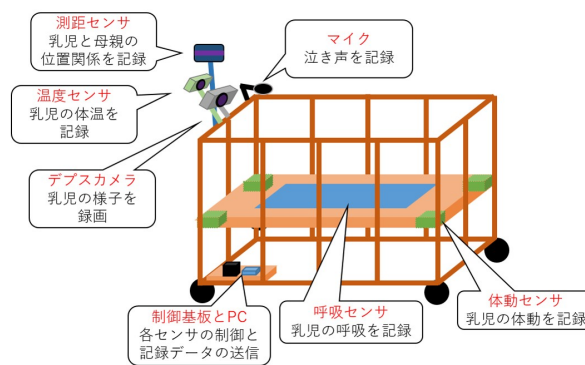


図1 ベビーベッド型デバイス概略図

器の装着や使用が不要であり、ベビーベッド自体は一般的なものを利用するため、使用に対する抵抗感やストレスは少ないと考えられる。こういった理由から、乳児の近くで、日常的に、育児の妨げとならないことを考慮したセンシングシステムが実現可能であると考え、ベビーベッドを選択した。

ただし、乳児の近くでセンシングを行うということは、乳児のすぐそばにセンサが存在することになるため、十分に安全性を考慮する必要がある。また、カメラやマイクによって映像や音声を取得する場合、一般家庭での運用を目的としている以上、使用者のプライバシーを侵害しないようなセンシングが求められる。これらを考慮し、このデバイスの開発においては、まずは乳児の状態推定に有用なデータを選別し、その上で最終的には推定のために必要最低限のデータを取得するセンサのみを取り付け、そのデータに対してもプリプロセッシングにより個人を特定するような情報が残らないようにする。このことは、長くデータを取り続けるが故のデータの量の肥大化への対策ともなりうる。このように、安全性とプライバシーを十分に考慮した、日常的に使うだけで乳児の状態推定が可能なベビーベッド型デバイスを開発し、育児の適切な支援を実現することを目標とする。

ベビーベッド型デバイスより取得するデータの選別にあたり、大阪大学大学院医学系研究科の先生方との相談をもとに、乳児の啼泣認識や原因推定に対し影響を与えると考えられる様々な要素を選別した。取得するデータに基づき設置するセンサと、各データの取得理由を記載した、ベビーベッド型デバイスの概略図を図1に示す。こちらのそれぞれのデータに対し、乳児の啼泣認識と原因推定において必要な情報であるかを検討していく。

予備実験として、ベビーベッド型デバイスの使用上での改善点を明確にするため、デバイスのプロトタイプを実際の家庭環境で使用してもらい、その際の意見や要望を求めた。また同時に、デバイスから安定して

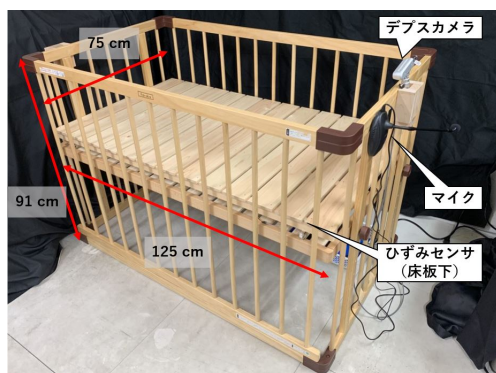


図2 ベビーベッド型デバイス外観

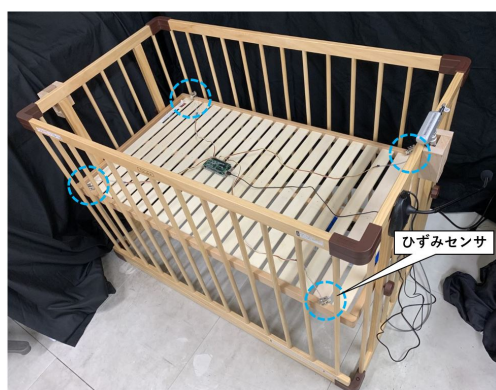


図3 ベビーベッド型デバイスの床板内部

ベビーベッド上の乳児の情報が取得できるかも確かめた。この実験は、生後1か月の乳児のいる家庭に、両親の同意のもとでベビーベッド型デバイスを設置し、通常のベビーベッドと同様に利用してもらうことで実施した。使用に関して、センシングされるデータの内容を伝え、デバイスに対する操作は必要ないことを伝えた。この時に利用したデバイスの外観を図2に、床板内のセンサ位置を図3に示した。今回の実験では、安全な設置が可能であったデプスカメラ (Intel RealSense Depth Camera D435)、マイク (FIFINE 製)、ひずみセンサ (ミネベアミツミ製) を設置した。

このデータ取得後に明らかになった主な改善点は、プライバシーに関するものであった。特に、マイクやカメラにより、取得されてほしくない個人情報まで含んでしまう可能性を指摘された。その意見に対して、マイクをスイッチの付いたものに変更し、カメラは深度映像のみを残すといった形で対処した。これらについてはベビーベッド型デバイスの実用時においても議論される問題ではあるが、映像や音声で啼泣認識や原因識別に不要と判断される可能性や、必要と判断された場合でも、最終的にはプリプロセッシングにより個人の特長ができない形で保存することを踏まえれば、対処可能であると考えられる。この他にも、センサのケーブルの位置が乳児の手に届きそうで不安といった、

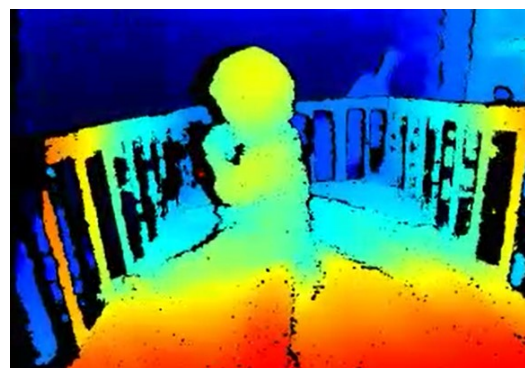


図4 ベッド上の様子

安全面への指摘を受けた。一方で、乳児の情報の取得に関しては、設置時には実運用における問題から取得ができなかったが、最終的には安定して取得できることを確認した。この結果から、デバイスの設計において、プライバシーや安全性の改善点が挙げられたものの、ベビーベッド上の乳児の情報を安定して取得できることが分かった。

4. 評価実験

予備実験に用いたベビーベッド型デバイス (図2) を、予備実験に基づいて改善した上で、雑音を含む一般家庭環境に設置し、乳児のセンシングを行った。本実験は、各データが乳児の啼泣認識において有用か、並びに複数の情報を組み合わせが性能向上において有用かを評価するために実施した。実験の対象となったのは、生後9か月の乳児1名であった。使用において、記録することが望ましくない音声が発せられる場合は、マイクの録音機能を切ってもよいことについて育児者に述べ、他に特別な指示はせず、普段通り利用してもらった。この時のベッド上の様子を、深度画像にて図4に示す。

得られたデータの内、それぞれに啼泣を含む3分間と2分間のデータ、合計5分間分のデータを抽出した。そのデータの内、安定して取得できた、音声と重心・重量変化の情報を利用して、啼泣認識の可能性について評価した。認識については、重心・重量変化の取得周波数を基に10Hzで行い、各フレーム毎に啼泣か非啼泣かを判断した。判定における乳児の啼泣区間は、協力者3名に啼泣音声を聴いてもらい、2人以上が啼泣だと判定した区間とした。啼泣区間の判定は、Anvil¹を用いて、0.1秒単位で行った。認識手法には、音声と重心・重量変化における周波数を利用し、閾値で判定する手法とSVMを用いた手法を検証した。評価には、検出漏れや誤検出の割合から導かれる統計量であるF値を用いた。

1: <https://www.anvil-software.org/>

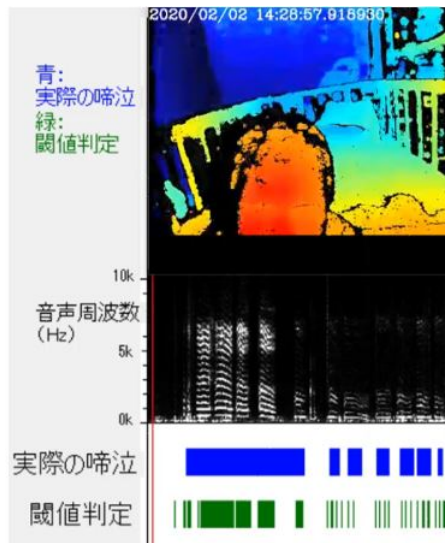


図5 啼泣認識の様子の一部

音声と重心・重量変化を周波数解析するにあたり、それぞれに対し短時間フーリエ変換を適用した。音声においては、先行研究により、乳児の啼泣の短時間処理の窓幅は、基本周波数を考慮して10~30msが望ましいとされており^[10]、取得した音声は44100 Hzであったことから、 $1024/44100 \approx 0.023(s)$ の関係により音声の窓関数の幅を1024とした。重心・重量変化の窓関数の幅は64とした。窓関数には、周波数分解能が良いハミング窓を選択した。

閾値で判定する手法では、短時間フーリエ変換によって得られた特定の周波数のスペクトルに対し、一定の閾値以上であった場合は啼泣と判定した。この時、もっともF値が大きくなるような周波数と閾値の組み合わせを求めることで評価した。音声は、44100 Hzのデータを幅1024の窓関数を用いて周波数スペクトルを求めたため、スペクトルを求めた周波数の間隔は43.07 Hzであった。0から0.1までの閾値を検証し、刻み幅は有効数字2桁で周囲のF値の変化が無くなった 1.0×10^{-5} とした。重心・重量変化は12 Hzのデータを幅64 Hzの窓関数を用いて周波数スペクトルを求めたため、スペクトルを求めた周波数の間隔は0.1875 Hzであった。0から1までの閾値を検証し、刻み幅は有効数字2桁で周囲のF値の変化が無くなった 1.0×10^{-4} とした。

重心・重量変化のみを用いた時のF値は、まず重量、重心x方向、重心y方向に対しそれぞれの値を求め、その後3つを組み合わせた時の値を求めた。重心と重量の値の組み合わせ方は、重量と重心x、y両方向の3要素を平等に扱うために、それぞれの周波数の最大値を用いて正規化した上で、3要素の内最大のスペクトル値を用いて判定するものとした。

表1 閾値判定のF値

特徴量	F値
音声のみ	0.55
重量のみ	0.49
重心x方向のみ	0.52
重心y方向のみ	0.52
重心・重量	0.51
音声と重量	0.53
音声と重心x方向	0.55
音声と重心y方向	0.56
音声と重心・重量	0.58

表2 SVMのF値

特徴量	F値
音声のみ	0.36
重量のみ	0.32
重心x方向のみ	0.48
重心y方向のみ	0.41
重心・重量	0.45
音声と重量	0.41
音声と重心x方向	0.44
音声と重心y方向	0.43
音声と重心・重量	0.46

音声と重心・重量を組み合わせた時の判定は、それぞれが閾値判定を行い、一方が啼泣と判断すれば啼泣とする時と、両方が啼泣と判断したときのみ啼泣とする時の両方を検証し、より大きい値のF値で評価した。

SVMを用いた手法では、啼泣区間と非啼泣区間を均等に分けた、5分割交差検証により評価した。SVMのカーネル関数はRBFとし、用いる特徴量は閾値の時と同様に短時間フーリエ変換によって得られた音声及び重心・重量変化の周波数のスペクトルとした。Cパラメータと γ パラメータの最適化を行い、その内最も大きなF値が得られた組み合わせを抽出し、そのF値によって評価した。

閾値判定により得られた特徴量毎のF値を、表1に示した。閾値を用いた時のF値は、音声及び重心・重量変化をそれぞれ単体で利用した時と比較して、音声と重心・重量変化を組み合わせた時の値が0.58と最も高くなった。この時の判定の内、1分間を抽出した時の様子を、図5に示した。

SVMにより得られた特徴量毎のF値を、表2に示した。SVMを用いた時のF値は、重心x方向のみを用いた時の値が0.48と最も高くなっている。しかしながら、閾値判定の値と比較すると、全体的に値が小さくなっていることが分かる。

5. 考察

実験結果より、閾値判定において音声と重心・重量変化を組み合わせた時に F 値が 0.58 と最も高くなり、音声と重心・重量変化の組み合わせにより啼泣認識の性能が向上できることが示唆されたと考える。またこれにより、複数のデータを組み合わせることが、認識精度の向上につながることも示唆され、複数のデータを取得し解析するベビーベッド型デバイスの有用性も示せたと考える。

SVM を用いた場合では、全体的に閾値判定の時より F 値が低くなっている。これは、今回評価の対象としたデータが 5 分間と短いものであり、さらにはその中で啼泣が占める割合が少なかったことから、サンプル数の不足により十分な学習ができなかったからだと考えられる。したがって、コロナウイルスの影響で被験者を募るのが難しい現状ではあるが、今後実験対象とする乳児や記録時間を増やしていくことで、認識性能が向上する可能性がある。

また、音声と重心・重量変化を組み合わせた場合の閾値判定における F 値 (0.58) は、啼泣の音声のみを用いた先行研究^[4]の値 0.597 よりも低い結果であった。その原因として、図 5 から分かるように、啼泣ではない区間に啼泣と判断されている箇所や、また啼泣中にも啼泣の判定が途切れている箇所があることが考えられる。誤った判定の原因としては、データの中には親を呼んでいるような乳児の甘え声と動作が含まれており、それが啼泣と判断されたため、誤った認識が生まれたと考えられる。これについては、音声と重心・重量変化により乳児の甘え声も取得できる可能性があるとも考えられ、原因推定の段階で識別可能となることが期待できる。また、今回対象にした乳児は生後 9 か月と年齢が高く、ベッドの上を自由に動き回ることができる年齢であったために、重心・重量変化の周波数特徴やマイクとの距離による音量が変わり、認識の漏れや認識の途切れが生じたと考えられる。育児者が最も不安を感じるのは、初めての子どもが生まれたばかりの頃であり、その年齢の乳児であればベッド上での動きは小さくなると考えられるため、啼泣時における体動の特徴が顕著になる可能性はある。しかしながら、動き回るような年齢であっても正確に啼泣を認識するシステムの実現も目指し、解析手法等の検討も進めていきたい。

検出漏れや誤検出への対策として、啼泣の持続時間を調査し、その情報から認識された啼泣を継続時間でフィルタリングする方法も挙げられる。また、今回は音声と重心・重量変化の情報のみを用いて評価を行ったが、ベビーベッド型デバイスではそれ以外にも乳児

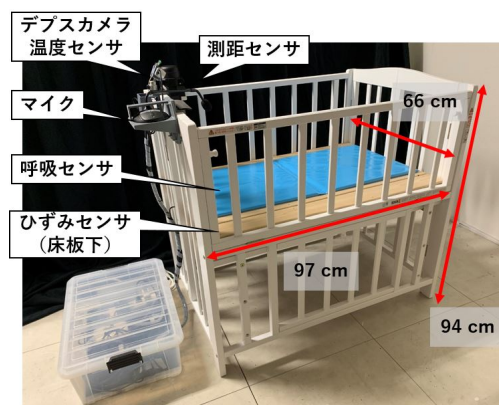


図 6 改良版ベビーベッド型デバイス外観

の体温や第三者の接近などの情報も取得することができるため、こうしたデータを特徴量に組み合わせる、あるいは音声や重心・重量変化と入れ替えることで、さらなる性能向上が期待できると考える。

6. 今後の展望

今回の実験を踏まえて、今後の実験のために、改良版のベビーベッド型デバイスを作成した。このデバイスの外観を図 6 に示した。用いているセンサは、デプスカメラ (Intel RealSense Depth Camera D435)、温度センサ (OMRON 非接触温度センサ D6T-44L-06)、距離センサ (SLAMTEC RPLiDAR S1)、マイク (FIFINE 製)、呼吸センサ (IBUKI ONE)、ひずみセンサ (ミネベアミツミ製) である。このデバイスは既に 2 組の一般家庭にて実際に利用してもらい、使用してもらった上で考慮すべき点を明確にした上で改良を行い、安定してデータの取得ができるようになったものである。今後はこのデバイスを用いて、より多くの乳児のデータを取得し、それを用いて乳児の啼泣認識や原因推定を試みる。

また、デバイスから取得したデータに関しては、啼泣の認識や識別以外にも様々な活用方法が期待できる。例えば、本稿で啼泣か否かの判定のみを用いて評価を行ったが、解析手法によっては、啼泣以前に何らかの兆候が読み取れる可能性もある。また啼泣以外にも、体重の変化や体温の推移などから、乳児の状態や成長度合いも確認できる可能性もある。そのような情報を自動的に記録するようなシステムは実現可能であるため、デバイス上に寝かせるだけで母子手帳を記録する支援ができ、育児者のストレスを軽減できる可能性がある。

7. おわりに

本稿では、日本の少子化問題の背景にある、乳児の啼泣原因が分からないといった不安から、子供を持と

うとしない人が増えている現状を改善するため、日常的に乳児の様々なデータを取得し、啼泣認識と原因推定により育児者を支援するための、ベビーベッドを用いたシステムを考案し、複数のデータを用いた啼泣認識の性能によってその有用性を評価した。実験の結果、音声と重心・重量変化を組み合わせることで判定した時に最も高い性能を示したため、複数の情報を利用した解析により、啼泣認識の精度が向上する可能性が示唆された。

謝辞

本研究は、大阪大学大学院医学系研究科の関係者の方々、及び実験にご協力いただいた方々から多大なるご助力を頂いたものであり、ここに深く感謝の意を表します。

なお本研究は、大阪大学研究倫理委員会の承認を得た。また、本研究の一部は文部科学省 Society5.0 実現化研究拠点支援事業「保健・予防医療プロジェクト生誕 1000 日見守り研究」の支援を受けた。

参考文献

- [1] 飯田, 他: 単胎初産婦の産後 1 か月までの育児不安; 森ノ宮医療大学紀要, **Vol.12**, pp.21–34 (2018).
- [2] 山崎, 篠原, 秋山, 市川, 尾島, 玉腰, 松浦, 山崎, 山縣: 乳幼児を持つ母親の育児不安と日常の育児相談相手との関連: 健やか親子 21 最終評価の全国調査より; 日本公衆衛生雑誌, **Vol.65**, No.7, pp.334–346 (2018).
- [3] Lavner, Y., Cohen, R., Ruinskiy, D., IJzerman, H.: Baby cry detection in domestic environment using deep learning; 2016 IEEE International Conference on the Science of Electrical Engineering, pp.1–5, (2016).
- [4] Yao, X., Micheletti, M., Johnson, M., de Barbaro, K.: Detection of Infant Crying in Real-World Home Environments Using Deep Learning; arXiv preprint arXiv:2005.07036v3 (2021).
- [5] Torres, R., Battaglino, D., Lepauloux, L.: Baby cry sound detection: A comparison of hand crafted features and deep learning approach; International Conference on Engineering Applications of Neural Networks, pp.168–179, (2017).
- [6] Arakawa, K.: Recognition of the cause of babies' cries from frequency analyses of their voice classification between hunger and sleepiness; Proc. ICA, pp.1713–1716 (2004).
- [7] Kurokawa, T., Miura, T., Yamashita, M., Sakai, T., Matsunaga, S.: Emotion-Cluster Classification of Infant Cries Using Sparse Representation; 2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, pp.1875–1878 (2018).
- [8] Parga, J. J., Lewin, S., Lewis, J., Montoya-Williams, D., Alwan, A., Shaul, B., et al: Defining and distinguishing infant behavioral states using acoustic cry analysis: is colic painful?; Pediatric research, **Vol.87**, No.3, pp.576–580 (2019).
- [9] Prasasti, A. L., Novamizanti, L., Razik, M. I.:

Identification of baby cry with Discrete Wavelet Transform, Mel Frequency Cepstral Coefficient and Principal Component Analysis; Journal of Physics: Conference Series, **Vol.1367**, No.1, p.012061 (2019).

- [10] Liu, L., Li, W., Wu, X., Zhou, B. X.: Infant cry language analysis and recognition: an experimental approach; IEEE/CAA Journal of Automatica Sinica, **Vol.6**, No.3, pp.778–788 (2019).