

SISTEM IDENTIFIKASI ARTI TANGISAN BAYI MENGGUNAKAN METODE MFCC, DWT DAN KNN PADA RASPBERRY PI

IDENTIFICATION SYSTEM OF BABY CRY USING MFCC, DWT AND KNN METHODS IN RASPBERRY PI

Retno Hadi Prasetyo¹, Iwan Iwut Tritoasmoro, S.T., M.T.², Ledy Novamizanti, S.Si., M.T.³

^{1,2,3}Prodi S1 Teknik Telekomunikasi, Fakultas Teknik Elektro, Telkom University

¹hadiprasetyo@student.telkomuniversity.ac.id, ²iwaniwut@telkomuniversity.ac.id,

³ledyaldn@telkomuniversity.ac.id

Abstrak

Penanganan kurang tepat terhadap bayi yang sedang menangis masih banyak dilakukan oleh banyak orang tua dan terkhusus yang baru memiliki anak. Hal ini dikarenakan orang tua tidak memahami arti dari tangisan bayi sehingga keinginan bayi tidak terpenuhi. *Dunstan Baby Language* (DBL) mengklasifikasikan bahasa bayi menjadi lima jenis kondisi yaitu EH (berarti sendawa), HEH (berarti tidak nyaman), NEH (berarti lapar), OWH (berarti mengantuk), EAIRH (berarti masuk angin). Pada Tugas Akhir ini telah dibuat sistem identifikasi arti tangisan bayi secara otomatis berdasarkan *speech processing* pada *raspberry pi*.

Rekaman suara tangisan bayi digunakan sebagai sinyal masukan kemudian diekstraksi ciri menggunakan *Mel-Frequency Cepstral Coefficient* (MFCC) dan *Discrete Wavelet Transform* (DWT). Hasil ekstraksi ciri diklasifikasikan menggunakan *K-Nearest Neighbor* (KNN). Keluaran sistem berupa 5 kelas suara kondisi bayi yaitu : tidak nyaman, lapar, masuk angin, sendawa atau mengantuk. Rancangan sistem identifikasi ini diimplementasikan pada *raspberry pi*.

Sistem ini menggunakan jumlah data latih sebanyak 50, jumlah data uji sebanyak 20 dan 4 suara bukan tangisan bayi. Parameter terbaik yang didapatkan dengan menggunakan jumlah *frame* sebanyak 512 data per *frame*, koefisien ciri MFCC sebanyak 39, jenis *Daubechies Wavelet* yaitu Db1 dan nilai *threshold* untuk *Euclidean distance* sebesar 30. Sistem dapat mengidentifikasi suara tangisan bayi dengan akurasi terbaik yaitu 90% dengan waktu komputasi selama 8,707 detik.

Kata Kunci: Tangisan Bayi, DBL, MFCC, DWT, KNN, Raspberry Pi.

Abstract

Inappropriate handling of babies who are crying is still done by many parents and especially those who have new children. This is because parents do not understand the meaning of the crying baby so that the baby's wishes are not fulfilled. Dunstan Baby Language (DBL) classifies baby language into five types of conditions namely EH (meaning belching), HEH (meaning uncomfortable), NEH (meaning hungry), OWH (meaning sleepy), EAIRH (meaning colds). In this Final Project a system has been made to identify the meaning of a baby's cry automatically based on speech processing on raspberrypi.

The baby's cry sound recording will be used as an input signal which is then extracted features using Mel-Frequency Cepstral Coefficient (MFCC) and Discrete Wavelet Transform (DWT) and the results of feature extraction will be classified using K-Nearest Neighbor (KNN). System output in the form of 5 classes of baby's sound conditions namely: discomfort, hungry, belly-pain, burp or sleepy. The design of this identification system will be implemented on raspberrypi.

This system uses the amount of training data of 50, the amount of test data of 20 and 4 sounds not crying baby. The best parameters obtained using 512 frames per frame, MFCC characteristic coefficient of 39, Daubechies Wavelet type is Db1 and the threshold value for Euclidean distance is 30. The system can identify the baby's crying sound with the best accuracy of 90% with computing time for 8.707 seconds.

Keywords: Baby's Cry, MFCC, DWT, KNN, Raspberry Pi.

1. Pendahuluan

Pada umumnya, seorang bayi akan menangis saat baru dilahirkan. Tangisan itu menunjukkan bahwa bayi terlahir dalam kondisi sehat. Karena belum adanya kemampuan untuk berbicara maka bayi hanya bisa menangis untuk mengekspresikan kondisi yang sedang dirasakan. Pada kondisi seperti ini, orang tua harus bisa memahami keinginan bayi melalui tangisannya agar tidak terjadi kesalahan penanganan terhadap sang bayi. Namun, pada kenyataannya masih banyak orang tua yang belum bisa memahami arti tangisan bayinya yang mengakibatkan bayi menangis secara terus menerus. Berdasarkan hasil penelitian dari *Dunstan Baby Language* yang dilakukan oleh musisi asal Australia bernama Priclilla Dunstan pada tahun 2006, menyatakan bahwa *Dunstan Baby Language* mengklasifikasikan tangisan bayi berdasarkan kebutuhannya menjadi lima jenis bahasa. Lima jenis bahasa yaitu EH (berarti sendawa), NEH (berarti lapar), HEH (berarti tidak nyaman), OWH (berarti mengantuk) dan EIARH (berarti masuk angin) [1].

Penelitian mengenai identifikasi tangisan bayi telah dilakukan oleh beberapa peneliti sebelumnya. Pada tahun 2018, peneliti dari Universitas Langlangbuana (UNLA) dan Institut Teknologi Nasional (ITENAS) telah melakukan penelitian tangisan bayi dengan menguji sebanyak 25 *Linier Predictive Coding* (LPC) dan *Euclidean Distance* sehingga mendapatkan akurasi sebesar 76% [2]. Pada tahun 2016, peneliti dari Institut Teknologi Sepuluh November juga melakukan penelitian tangisan bayi dengan menguji sebanyak 139 sampel suara tangisan bayi dengan 5 suara tangisan bayi tiap kondisi menggunakan metode *Mel-Frequency Cepstral Coefficient* (MFCC) dan *K-Nearest Neighbor* (KNN) sehingga mendapatkan hasil akurasi sebesar 79,95% [3].

Oleh karena itu pada Tugas Akhir ini akan membuat rancangan sistem identifikasi tangisan bayi pada *raspberry pi* menggunakan metode ekstraksi ciri yaitu *Mel-Frequency Cepstral Coefficient* (MFCC), *Discrete Wavelet Transform* (DWT) dan metode klasifikasi yaitu *K-Nearest Neighbor* (KNN). Keluaran dari sistem identifikasi ini berupa suara yang menjelaskan kondisi bayi. Tugas Akhir ini menggunakan metode MFCC karena dalam proses pengenalan suara mirip dengan telinga manusia dan memiliki tingkat akurasi sebesar 99.78% [13] [14]. Selain MFCC, fitur ekstraksi yang digunakan adalah DWT yang dapat mengurangi noise dan mengkompresi sinyal tanpa degradasi yang berlebih [15] [16].

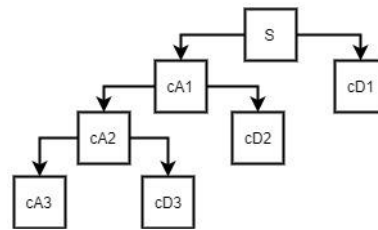
2. Dasar Teori

Berdasarkan *Dunstan Baby Language* (DBL), tangisan bayi diklasifikasikan menjadi 5 kondisi. Kelima kondisi tersebut adalah lapar (NEH), mengantuk (OWH), bersendawa (EH), tidak nyaman (HEH), masuk angin (EAIARH). *Dunstan Baby Language* (DBL) diperkenalkan oleh musisi asal Australia bernama Priclilla Dunstan pada tahun 2006. Dia menyatakan bahwa kelima kondisi tersebut berlaku sama bagi semua bayi dan tidak memandang ras, suku dan bahasa [1] [4].

Discrete Wavelet Transform (DWT) adalah salah satu metode ekstraksi ciri dengan cara menganalisis sinyal. *Transform wavelet* dikembangkan sebagai alternatif *short time fourier transform* (STFT). DWT juga digunakan untuk mentransformasikan isyarat dari domain waktu ke domain frekuensi. Perhitungan DWT secara umum dapat dilihat pada persamaan 2.1 [17].

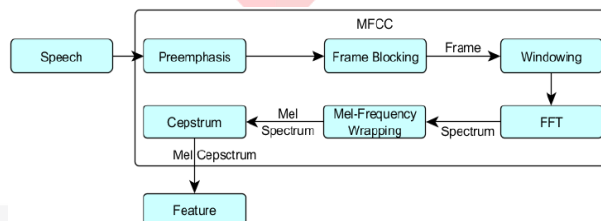
$$x[n] = \sum_{k=-\infty}^{\infty} (y_{high}[k] \cdot g[-n + 2k]) + (y_{low}[k] \cdot h[-n + 2k]) \quad (1)$$

dimana a dan b adalah bilangan real, $*$ adalah notasi konjugate kompleks dan (t) adalah fungsi *wavelet*. DWT memiliki dua filter yaitu Low Pass Filter (LPF) dan *High Pass Filter* (HPF). Dua filter ini digunakan untuk melakukan proses dekomposisi sinyal asli [8]. HPF digunakan untuk filterisasi dan analisis sinyal yang berfrekuensi tinggi sedangkan LPF digunakan untuk filterisasi dan analisis sinyal yang berfrekuensi rendah. Setelah filterisasi akan dihasilkan sebuah resolusi yang digunakan untuk analisis terhadap frekuensi. Proses analisis terhadap frekuensi ini disebut dengan *multi-resolution analysis* [9]. Pada Gambar 1 merupakan proses *filtering* pada DWT, dimana S adalah sinyal awal yang akan di dekomposisi, $cA1$ merupakan sinyal LPF, $cD1$ merupakan sinyal HPF. Proses *filtering* dilakukan sampai menemukan level yang diinginkan.



Gambar 1. Level DWT [9].

Mel Frequency Cepstrum Coefficient (MFCC) adalah metode ekstraksi fitur sinyal suara. Ekstraksi fitur adalah proses penentuan nilai atau vektor yang dapat digunakan sebagai objek atau identitas individu [8]. Secara umum, MFCC digunakan sebagai parameter fitur fonologis dalam aplikasi pengenalan speaker [2]. Analisis suara dilakukan setelah mengambil input melalui mikrofon dari pengguna. Desain sistem melibatkan manipulasi sinyal audio input. Pada level yang berbeda, operasi yang berbeda dilakukan pada sinyal input seperti *pre-emphasis*, *framing*, *windowing*, analisis Mel Cepstrum dan pencocokan dari suara tangisan bayi. Gambar 2 merupakan proses LPC untuk mendapatkan ekstraksi fitur dari MFCC.



Gambar 2. Proses LPC [8].

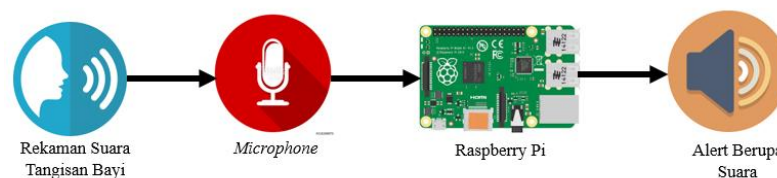
K-Nearest Neighbor (KNN) merupakan metode yang digunakan untuk klasifikasi data uji berdasarkan atribut dan sampel data *training* [10]. Diberikan data uji yang digunakan sebagai masukan sistem, kemudian pada data *training* akan dicari nilai *K* yang dekat dengan nilai *K* data uji. Ketika terdapat beberapa nilai *K* data *training* yang dekat dengan nilai data uji maka akan dipilih nilai *K* data *training* yang memiliki jarak terdekat dengan data uji (hampir sama dengan prinsip kerja algoritma djijkstra). Untuk menentukan dekat atau jauhnya jarak antara nilai *training* dengan nilai data maka dapat menggunakan persamaan *euclidean distance* [3].

$$d(p, q) = \sqrt{\sum_{i=1}^N (q_i - p_i)^2} \quad (2)$$

dimana $d(p, q)$ adalah jarak *Euclidean Distance* antara vektor p dan vektor q , N adalah jumlah data pada kedua vektor, q_i adalah data ke- i pada vektor q , p_i adalah data ke- i pada vektor p .

3. Perancangan

Tugas Akhir ini bertujuan untuk merancang sistem identifikasi arti tangisan bayi pada *raspberry pi*. Desain sistem merupakan rancangan cara kerja sistem secara keseluruhan. Gambar 3 adalah desain sistem pada Tugas Akhir ini.

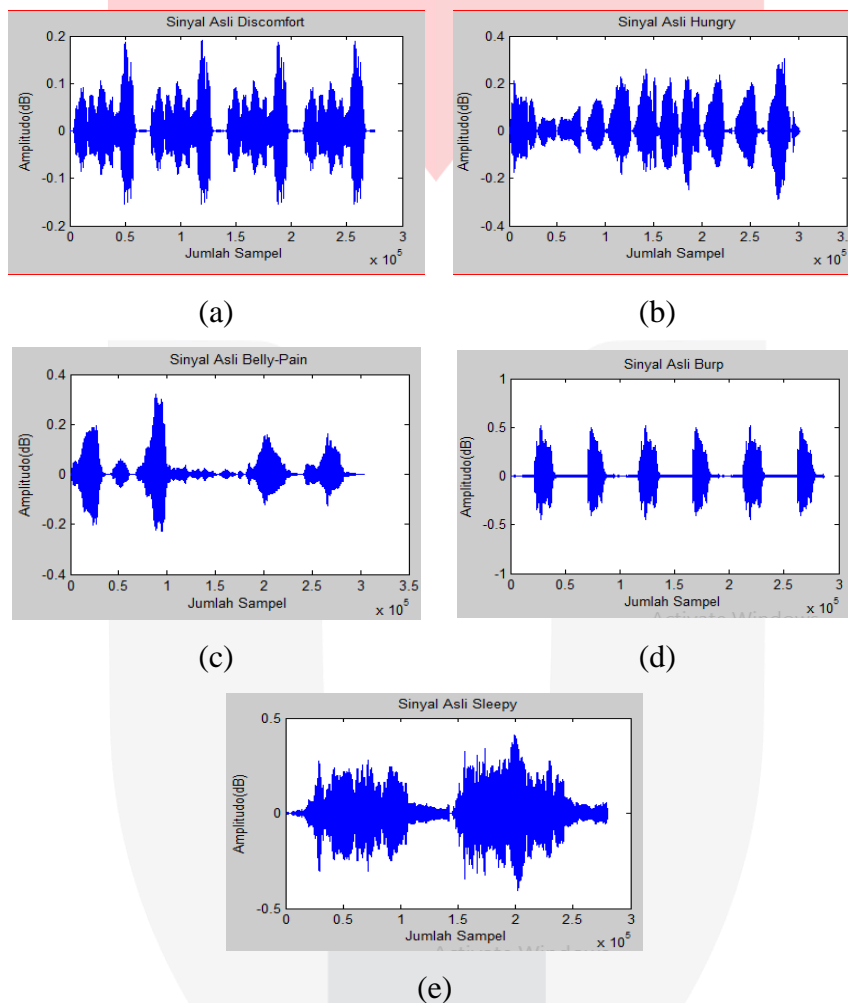


Gambar 3. Desain sistem.

Berdasarkan Gambar 3 dapat dijelaskan bahwa sinyal masukan berasal dari suara tangisan bayi yang direkam dengan menggunakan alat perekam suara. Kemudian suara hasil rekaman diputar dan didekatkan pada *microphone* yang terhubung ke *raspberry pi*. Lalu suara yang masuk akan diekstraksi ciri dengan metode *Mel-Frekuensi Cepstral Coefficient (MFCC)*, *Discrete Wavelet Transform (DWT)* dan akan diklasifikasikan dengan metode *K-Nearest Neighbor (KNN)*. Hasil dari

klasifikasi ini akan mengaktifkan *alert* berupa suara yang menunjukkan 5 kelas kondisi bayi yaitu tidak nyaman, lapar, masuk angin, sendawa atau mengantuk.

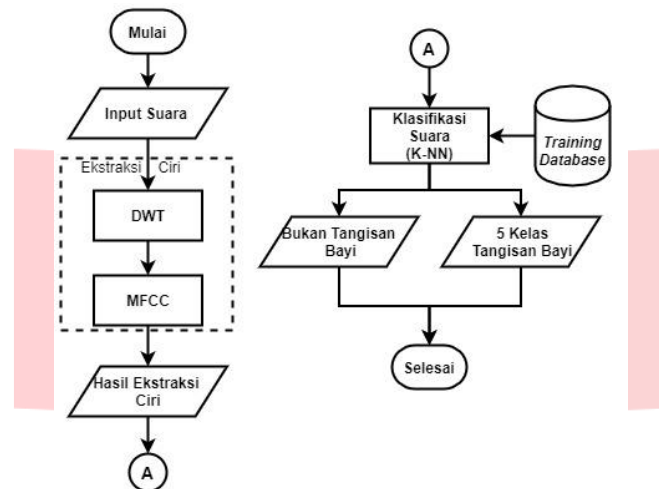
Pada Tugas Akhir ini, suara rekaman berupa *file wav* yang memiliki rata-rata durasi selama 7 detik. Alat yang digunakan untuk proses perekaman yaitu sebuah *smartphone* yang terdapat fitur perekam suara. Suara rekaman berasal dari 60 bayi dan tiap bayi rata-rata direkam sebanyak dua kali. Perekaman dilakukan dengan cara mendekatkan *smartphone* ke bayi dengan jarak 7 cm untuk mengurangi noise yang masuk. Data yang diambil diverifikasi oleh perawat yang jaga pada tempat yang sedang dilakukan pengambilan data. Pada proses *training* sistem menggunakan rekaman suara sebanyak 50 suara. Pada proses pengujian menggunakan 20 suara dan 4 suara bukan tangisan bayi yang digunakan untuk pengujian dalam kasus deteksi bukan suara tangisan bayi. 4 suara bukan tangisan bayi tersebut yaitu 1 suara tangisan orang dewasa, 1 suara hujan, 1 suara kendaraan dan 1 suara musik. Jadi jumlah keseluruhan suara yang digunakan pada penelitian ini sebanyak 74 suara.



Gambar 4. Sinyal tangisan bayi tiap kelas. (a) *discomfort*, (b) *hungry*, (c) *belly-pain*, (d) *burp*, (e) *sleepy*.

Gambar 5 adalah diagram alir cara kerja sistem secara keseluruhan. Diagram alir cara kerja sistem pada Gambar 5 bertujuan untuk memperjelas alur pengujian sistem secara keseluruhan. Berdasarkan Gambar 5 bahwa sinyal masukan berupa rekaman suara tangisan bayi yang didekatkan pada *microphone* lalu diekstraksi ciri dengan metode MFCC dan DWT. Setelah itu, hasil ekstraksi ciri itu dicocokkan dengan data hasil *training* dengan metode KNN. Apabila sinyal masukan cocok atau tidak sesuai dengan data suara hasil *training* maka akan mengaktifkan *alert* berupa suara yang menunjukkan 5 kelas kondisi bayi yaitu tidak nyaman, lapar, masuk angin, sendawa, atau mengantuk. Apabila sinyal masukan tidak cocok atau tidak sesuai dengan data suara hasil *training*

maka sistem akan otomatis berhenti melakukan proses dan kembali ke proses awal yaitu ke rekaman suara selanjutnya.



Gambar 5. Diagram alir cara kerja sistem.

4. Hasil dan Pembahasan

Pada pengujian pertama, sistem diuji dengan jumlah koefisien ciri MFCC sebesar 24, 30, dan 39. Koefisien ciri MFCC pada Tabel 1 disimbolkan dengan huruf C besar. Pada kasus ini, jumlah koefisien ciri MFCC sebesar 39 memiliki kualitas deteksi yang optimal daripada jumlah koefisien yang lain dengan tingkat akurasi 90% dan dengan waktu komputasi selama 8,8032 detik. Sedangkan untuk jumlah koefisien MFCC 24 dan 30, memiliki waktu komputasi yang singkat daripada koefisien MFCC 39. Semakin banyak jumlah koefisien MFCC, maka semakin teliti deteksi, namun semakin lama pula waktu komputasi. Berdasarkan hasil pengujian pada Tabel 1 dapat disimpulkan bahwa parameter yang paling optimal pada sistem ini adalah dengan frekuensi *sampling* 44100 Hz, jumlah *frame* 512, dan jumlah koefisien ciri MFCC sebesar 39 dengan waktu komputasi rata-rata 8,8032 detik.

Tabel 1. Hasil Pengujian terhadap Jumlah Koefisien Ciri MFCC.

Parameter Jumlah Frame = 512 dan Db 1							
C	Akurasi (%)	Waktu Komputasi (s)	Persentase Benar Tiap Kelas (%)				
			I	II	III	IV	V
24	48	4,1214	40	40	30	60	70
30	76	7,588	80	60	90	100	50
39	90	8,8032	90	90	90	90	90

Pada pengujian kedua, sistem diuji dengan panjang *frame* 128, 256, 512 dan 1024. Parameter yang digunakan pada pengujian ini adalah frekuensi *sampling* sebesar 44100 Hz, dan koefisien ciri MFCC sebanyak 39. Tabel 2 menunjukkan hasil dari pengujian terhadap jumlah *frame*. Pada kasus ini, tingkat akurasi bersifat fluktuatif dan pengujian dengan jumlah *frame* 512 memiliki tingkat akurasi paling tinggi yaitu sebesar 90%. Hal ini dikarenakan pada saat pengidentifikasian sinyal per *frame* terdapat *frame* yang tidak teridentifikasi secara optimal. Untuk waktu komputasi terukur selama 8,8032 detik. Dari kolom jumlah *frame* dan waktu komputasi pada Tabel 2 terlihat bahwa jika jumlah *frame* dikalikan dua maka waktu komputasi juga akan berlipat dua kali. Hal ini dikarenakan semakin banyak *frame* berarti semakin banyak segmentasi sehingga proses identifikasi semakin detail dan tingkat akurasi akan semakin tinggi. Namun banyaknya segmentasi berakibat pada lamanya waktu komputasi.

Tabel 2. Hasil Pengujian terhadap Jumlah *Frame*.

Parameter C = 39 dan Db 1							
Jumlah Frame	Akurasi (%)	Waktu Komputasi (s)	Persentase Benar Tiap Kelas (%)				
			I	II	III	IV	V
128	66	3,085	60	70	70	80	50
256	76	4,8042	70	80	80	90	60
512	90	8,8032	90	90	90	90	90
1024	80	14,613	80	70	80	100	70

Pada pengujian ketiga, sistem diuji dengan jenis *Daubechies Wavelet* yaitu Db1, Db2, dan Db8. Parameter yang digunakan pada pengujian ini adalah frekuensi sampling sebesar 44100 Hz, jumlah *frame* yang digunakan adalah 512 dan koefisien ciri MFCC sebesar 39. Tabel 3 menunjukkan hasil pengujian terhadap pengaruh *Daubechies Wavelet*. Berdasarkan pengamatan Tabel 3, jenis *Daubechies Wavelet* yaitu Db1 memiliki tingkat akurasi paling tinggi dan waktu komputasi pada masing-masing jenis *Daubechies Wavelet* tidak memiliki perbedaan yang begitu besar. Tidak seperti yang terjadi pada kasus pengujian pengaruh jumlah koefisien ciri MFCC dan pengaruh jumlah *frame* yang memiliki perbedaan waktu komputasi yang besar. Hal ini karena Db1 merupakan jenis *Daubechies Wavelet* yang sederhana yaitu hanya satu kali dekomposisi sehingga sistem dapat dengan mudah melakukan ekstraksi ciri. Dari pembahasan diatas dapat disimpulkan bahwa jenis *Daubechies Wavelet* yang baik untuk parameter yaitu Db1.

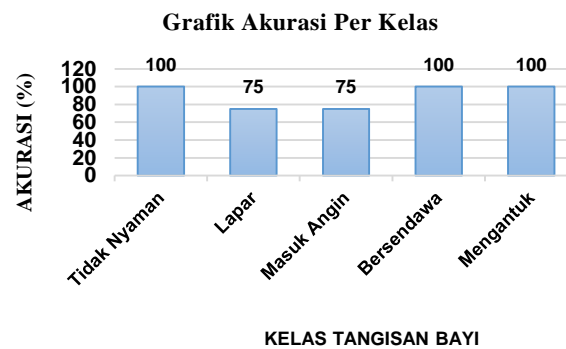
Tabel 3. Pengaruh *Daubechies Wavelet*.

Parameter Jumlah Frame = 512 dan C = 39							
Daubechies Wavelet	Akurasi (%)	Waktu Komputasi (s)	Parameter C = 8 dan Db 1				
			I	II	III	IV	V
Db1	90	8,8032	90	90	90	90	90
Db2	84	8,5304	70	80	100	100	70
Db8	66	7,552	50	50	60	100	70

Pada pengujian keempat dilakukan pengujian *database* terhadap data uji. Data uji berisi suara yang berbeda dengan suara yang digunakan untuk proses *training*. Data uji yang digunakan berjumlah 20 suara. Pada pengujian ini juga mempertimbangkan nilai K. K merupakan jumlah koefisien ketetangaan pada metode klasifikasi KNN. Berdasarkan Tabel 4 diperoleh tingkat akurasi terbaik yaitu pada saat nilai K = 5. Hal ini karena semakin besar dimensi atau kelas maka nilai K juga harus semakin besar sehingga memiliki lebih banyak pilihan pada saat pengklasifikasian. Gambar 6 menunjukkan bahwa kelas tangisan bayi yang masih terdapat kesalahan identifikasi yaitu kelas lapar dan kelas masuk angin. Setelah dilakukan perhitungan diperoleh tingkat akurasi sistem sebesar 90% dan rata-rata waktu komputasi selama 8,707 detik yang artinya dalam sekali siklus kerja, sistem membutuhkan waktu 8,707 detik.

Tabel 4. Pengujian Data Uji.

Jumlah Frame	Jumlah Koefisien Ciri	Daubechies Wavelet	Nilai K	Akurasi (%)	Waktu Komputasi (s)
512	39	Db1	K = 1	10	1,672
			K = 3	50	3,333
			K = 5	90	8,707



Gambar 6. Grafik Akurasi Per Kelas.

Pada pengujian kelima dilakukan pengujian *database* terhadap suara bukan tangisan bayi. Suara bukan tangisan bayi merupakan suara yang selain suara tangisan bayi. Pada pengujian ini menggunakan 4 suara bukan tangisan bayi dengan rincian 1 suara musik, 1 suara orang dewasa, 1 suara hujan, dan 1 suara kendaraan. Pada Tabel 5 menunjukkan hasil pengujian dengan suara bukan suara tangisan bayi.

Tabel 5. Pengujian Terhadap Suara Bukan Tangisan Bayi.

No.	Nama File	Teridentifikasi	Hasil Identifikasi	Waktu Komputasi (s)	Nilai Euclidean Distance
1	Suara Musik	Bukan Suara Bayi	Benar	8,95	32,338395
2	Suara Orang Dewasa	Bukan Suara Bayi	Benar	9,23	32,185333
3	Suara Hujan	Bukan Suara Bayi	Benar	8,82	36,291607
4	Suara Kendaraan	Bukan Suara Bayi	Benar	8,91	48,620768

Berdasarkan Tabel 5, semua suara bukan tangisan bayi yang diujikan tidak teridentifikasi sebagai suara tangisan bayi, melainkan dianggap *noise*. Hal itu dapat diketahui melalui nilai *Euclidean distance* dari suara bukan tangisan bayi, semuanya memiliki nilai yang tidak memenuhi nilai *threshold* yang telah ditentukan yaitu 30. Sehingga sistem ini dapat mengidentifikasi bahwa suara yang terekam merupakan suara tangisan bayi atau bukan.

5. Kesimpulan

Pada penelitian ini, telah dirancang suatu sistem deteksi suara tangisan bayi menggunakan metode DWT, MFCC, dan KNN. Sistem ini menggunakan jumlah data latih sebanyak 50 dan jumlah data uji sebanyak 20. Suara bukan tangisan bayi juga digunakan dalam menguji keakuratan sistem dengan menggunakan suara motor, suara mesin, suara orang dewasa dan suara hujan. Berdasarkan analisis yang telah dilakukan dapat disimpulkan bahwa sistem dapat mendeteksi suara tangisan bayi dengan menggunakan parameter terbaik yaitu jumlah *frame* sebanyak 512 data per *frame*, koefisien ciri MFCC sebanyak 39, dan jenis Daubechies Wavelet yaitu Db1. Hasil akurasi terbaik yang diperoleh sebesar 90% dengan waktu komputasi selama 8,707 detik. Berikut ini adalah beberapa saran yang diberikan untuk pengembangan penelitian ini:

1. Menggunakan *microphone* yang memiliki sensitivitas tinggi dan mencoba menggunakan *raspberrypi* versi terbaru yaitu *raspberrypi v4*.
2. Mencoba menggunakan metode klasifikasi yang lain misalnya LDA atau PCA.
3. Menambahkan keluaran berupa tampilan teks pada LCD dan mencoba mengintegrasikan dengan perangkat lain misalnya aplikasi android pada smartphone atau website.

Daftar Pustaka

- [1] P. Dunstan, "Open Up and Discover Your Baby's Language", 2012. [Online]. Available: [www.babytaal.nl/media/PDF/ComprehensiveBooklet\(2\).pdf](http://www.babytaal.nl/media/PDF/ComprehensiveBooklet(2).pdf) [25-Oct-2019]

- [2] T.N. Bandung and L. Bandung , “Identifikasi Suara Tangisan Bayi Menggunakan Metode LPC dan Euclidean Distance”, *Elkomika*, vol.6, No.1, pp 153-164, 2018.
- [3] W. S. Limantoro, C. Fatichah, and L. Yuhana, “Rancang Bangun Aplikasi Pendeteksi Suara Tangisan Bayi”, *Jurnal Teknik ITS*, Vol. 5, No. 2, ISSN: 2337-3539, 2016.
- [4] Anggaraeni D., et al, “The Implementation of Speech Recognition Using Mel-Frequency Cepstrum Coefficients (MFCC) and Support Vector Machine (SVM) Method Based On Python To Control Robot Arm”, *AASEC*, 2017.
- [5] A.L.Prasasti, L. Novamizanti, M.I. Razik, “Identification of Baby Cry with Discrete Wavelet Transform, Mel Frequency Cepstral Coefficient and Principal Component Analysis”, *Journal of Physics: Conference Series*, Vol 1367, 2019.
- [6] P.D.Wananda, L.Novamizanti, R.D. Atmaja, “Sistem Deteksi Cacat Kayu dengan Metode Deteksi Tepi SUSAN dan Ekstraksi Ciri Statistik”.*ELKOMIKA:Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi dan Elektronika*, Vol 6, No. 1, 2018.
- [7] Ecadio, “Speaker 3 Watt 4 Ohm”,2018.[Online]. Available: <https://ecadio.com/jual-speaker-4-ohm-3-watt>. [Accessed on 13 Juli 2020, 15:00:00 WIB].
- [8] O.W. Yuliantari, Risky, Risanuri Hidayat, “Ekstraksi Ciri dan Pengenalan Tutur Vokal Bahasa Indonesia Menggunakan Metode DWT dan DTW Secara Realtime”. *Proceeding SNST 7th*, Fakultas Teknik, Universitas Wahid Hasyim Semarang, 2017.
- [9] J. Givary, R.P. Astuti, L. Novamizanti, “Analysis Effect of Discrete Wavelet Transform in Multi Carrier Code Division Multiple Access”, *IEEE Asia Pacific Conference on Wireless and Mobile*.2016.
- [10] Lazada, “Mic Microphone Mini 3.5mm”,2020.[Online]. Available: <https://www.lazada.co.id/products/mic-microphone-mini-35mm-i144854444.html>. [Accessed on 13 Juli 2020, 15:00:00 WIB].
- [11] *Raspberry pi*, “Schematic *Raspberry pi* Model 3 B+”,2019.[Online]. Available:<https://www.raspberrypi.org/documentation/hardware/raspberrypi/schematics/README.md>. [Accessed on 30 Oktober 2019, 13:00:00 WIB].
- [12] Me Hobby, “Discover The Prototyping Around The *Raspberry pi* Nano Computer”,2019.[Online]. Available: <https://shop.mchobby.be/en/raspberry-s-kits/385-raspberry-pi-3-b-plus-hack-starter-pack-raspberry-included-3232100003859.html>. [Accessed on 30 Oktober 2019, 14:00:00 WIB].
- [13] Sunitha C., et al, “Speaker Recognition using MFCC and Improved Weighted Vector Quantization Algorithm”, *IJET*, vol 7, No. 5, 2015.
- [14] Gulzar Taabish, Singh Anand, Sharma Sandeep, “Comparative Analysis of LPCC, MFCC, and BFCC for the Recognition of Hindi Words using Artificial Neural Networks”, *IJCA*, vol 101, No. 12, 2014.
- [15] Cutajar Michelle, et al, “Comparative Study of Automatic Speech Recognition Techniques”, *IET*, vol 7, Issue 1, 2013.
- [16] V. Janse Pooja, et al, “A Comparative Study between MFCC and DWT Feature Extraction Techniques”, *IJERT*, vol 3, Issue 1, 2014.
- [17] Devi Anita, Misal Abhishek, Dr. Shina G.R., "Performance analysis of DWT at different levels for feature extraction of PCG signals", *ICMiCR*, 2013.
- [18] Janse Pooja V., Magre Smita B., Kruzekar Pratik K., "A Comparative Study between MFCC and DWT Feature Extraction Technique", *IJERT*, vol.3, Issue 1, January 2014.
- [19] Setiawan Angga, Hidayatno Achmad and Isnanto Rizal R.,”Aplikasi Pengenalan Ucapan dengan Ekstraksi Mel-Frequency Cepstrum Coefficients (MFCC) Melalui Jaringan Syaraf Tiruan (JST) untuk Mengoperasikan Cursor Komputer”, Maret, 2011.
- [20] Kita Informatika, “Hitung Manual Algoritma k-Nearest Neighbor (k-NN) Menggunakan Dataset Sederhana”,2019.[Online]. Available:<http://www.kitainformatika.com/2019/10/hitung-manual-algoritma-knearest.html>. [Accessed on 13 Juli 2020, 14:00:00 WIB].