

Sistem Pengenalan Intensitas Emosi Marah melalui Ucapan menggunakan Ekstraksi *Wavelet-Based Frequency Cepstral Coefficients* dan Algoritma *K-Nearest Neighbor* berbasis Raspberry Pi 4

M. Ihsan An-Nashir¹, Barlian Henryanu Prasetyo²

Program Studi Teknik Komputer, Fakultas Ilmu Komputer, Universitas Brawijaya
Email: ¹ihsanannashir@student.ub.ac.id, ²barlian@ub.ac.id

Abstrak

Emosi merupakan perasaan intens yang ditujukan kepada seseorang ataupun sesuatu. Salah satu cara yang dapat dilakukan oleh manusia dalam mengekspresikan emosi marah yang dapat dikenali adalah melalui ucapan. *Speech emotion recognition* merupakan suatu teknologi yang dapat digunakan untuk mengidentifikasi emosi dari suara orang yang sedang berbicara. Pada penelitian ini, dilakukan proses *speech emotion recognition* menggunakan metode ekstraksi WFCC (*Wavelet-based Frequency Cepstral Coefficients*), dimana metode ini menggunakan transformasi wavelet dalam ekstraksinya sehingga memiliki kemampuan untuk memisahkan berbagai macam variasi frekuensi pada waktu yang beragam. Selain itu, penelitian ini juga menguji kemampuan algoritma *K-Nearest Neighbor* dalam mengklasifikasikan intensitas emosi marah dari sinyal suara. Penelitian ini dilakukan menggunakan perangkat keras Raspberry Pi 4. Penelitian ini menghasilkan kesimpulan bahwa metode ekstraksi WFCC cukup efektif dalam mendeteksi intensitas emosi marah *rage* atau tinggi (HIGH) dengan akurasi sebesar 66,67%.

Kata kunci: Suara, Emosi Marah, *Wavelet-based Frequency Cepstral Coefficients*, Transformasi Wavelet, *K-Nearest Neighbor*, *Speech Recognition*.

Abstract

Emotion is an intense feeling directed towards someone or something. One way that humans can express angry emotions that can be recognized is through speech. Speech emotion recognition is a technology that can be used to identify emotions from the sound of someone speaking. In this research, a speech emotion recognition process was carried out using the WFCC (Wavelet-based Frequency Cepstral Coefficients) method, which uses wavelet transformation in its extraction and has the ability to separate various frequency variations at various times. In addition, this study also tested the ability of the K-Nearest Neighbor algorithm in classifying the intensity of angry emotions from sound signals. This study was conducted using a Raspberry Pi 4. This study concluded that the WFCC extraction method is quite effective in detecting high intensity angry emotions with an accuracy of 66.67%.

Keywords: *Speech, Angry Emotion, Wavelet-based Frequency Cepstral Coefficients, Wavelet Transform, K-Nearest Neighbor, Speech Recognition.*

1. PENDAHULUAN

Emosi merupakan perasaan intens yang ditujukan kepada seseorang ataupun sesuatu. Selain itu, emosi juga dapat diartikan sebagai reaksi yang timbul akibat perbuatan seseorang ataupun kondisi tertentu (Rahmawanthi, et.al., 2019). Emosi terdiri dari berbagai macam jenis, dan salah satu diantaranya adalah emosi marah. Marah merupakan suatu emosi yang secara fisik mengakibatkan peningkatan tekanan darah,

denyut jantung, serta tingkat adrenalin dan noradrenalin (Davis, 2021). Layaknya jenis emosi yang lain, marah pun memiliki tingkatan intensitasnya sendiri. Intensitas emosi marah dibagi menjadi tiga, yakni *Annoyance* (rendah), *Anger* (sedang), dan *Rage* (tinggi) (Gumelar et.al., 2020). Salah satu cara yang dapat dilakukan oleh manusia dalam mengekspresikan emosi marah yang dapat dikenali adalah melalui ucapan.

Ucapan merupakan salah satu bentuk

interaksi antar manusia dalam bentuk suara yang terdiri dari kata yang diucapkan dengan berbagai cara. Suara yang keluar dari ucapan-ucapan tersebut memiliki karakteristik dan frekuensi masing-masing, dimana karakteristik dan frekuensi tersebut dapat dijadikan acuan untuk mengidentifikasi jenis ucapan dan emosi apa yang dikeluarkan. Deteksi emosi manusia dapat dilakukan menggunakan teknologi seperti *speech recognition*.

Speech Recognition merupakan bidang yang relatif baru dimana mesin diproyeksikan untuk dapat memahami suara yang keluar dari mulut manusia. Menurut Yahya (2012), terdapat hubungan yang erat antara cara berbicara dan emosi dari pembicara, termasuk dari segi nada, energi, artikulasi, dan bentuk *spectral*. Hal tersebut dapat membentuk suatu ciri yang merepresentasikan suatu bentuk emosi dalam berbicara. Berdasarkan hal tersebut, pada penelitian ini peneliti tertarik untuk membuat sistem yang dapat mendeteksi kondisi emosional seseorang menggunakan *speech recognition* yang dimana lingkup penelitian ini adalah untuk mendeteksi emosi marah berdasarkan tiga tingkatan intensitasnya, yakni *Annoyance* (rendah), *Anger* (sedang), dan *Rage* (tinggi).

Pada penelitian ini peneliti tertarik untuk merancang sistem yang berfungsi untuk mengenali intensitas emosi marah seseorang melalui suara yang dikeluarkan, dengan cakupan intensitas emosi rendah, sedang, dan tinggi melalui pemrosesan suara. Pada penelitian ini peneliti mengusulkan Transformasi Wavelet sebagai metode untuk melakukan ekstraksi fitur. Transformasi Wavelet digunakan untuk melakukan analisis terhadap sinyal-sinyal dengan kandungan frekuensi yang bervariasi terhadap waktu (non-stationer). Selain itu, Transformasi Wavelet juga memiliki kemampuan untuk memisahkan berbagai macam karakteristik pada skala yang beragam (Anant dan Dowla, 1997). Peneliti juga menggunakan Raspberry Pi 4 yang bertindak sebagai mikrokomputer, dan untuk tahapan dari sistemnya sendiri terdiri dari *preprocessing*, *modelling*, dan *classification*. Pada tahap *preprocessing*, dataset suara diolah dan diekstraksi menggunakan metode *Mel-frequency cepstral coefficients* (MFCC) namun tidak menggunakan transformasi *Fast Fourier transform* (FFT), melainkan menggunakan Transformasi Wavelet, sehingga dikembangkan sebagai metode *Wavelet-Based Frequency Cepstral Coefficients* (WFCC). Selanjutnya,

Hasil dari *preprocessing* akan masuk ke tahap *modelling* dan selanjutnya akan dilakukan tahap *classification* dimana kedua tahapan tersebut dilakukan menggunakan algoritma *K-Nearest Neighbor* (KNN).

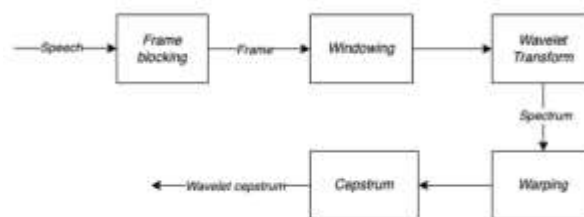
2. DASAR TEORI

2.1 Emosi Marah

Marah merupakan suatu emosi yang secara fisik mengakibatkan peningkatan tekanan darah, denyut jantung, serta tingkat adrenalin dan noradrenalin (Davis, 2021). Pada emosi marah, terdapat beberapa tingkatan intensitas yang terbagi menjadi *Annoyance* (rendah), *Anger* (sedang), dan *Rage* (tinggi) (Gumelar et al., 2020).

2.2 Wavelet-Based Frequency Cepstral Coefficients (WFCC)

Wavelet-Based Frequency Cepstral Coefficients (WFCC) merupakan metode yang digunakan dalam melakukan ekstraksi fitur dari sinyal suara pada penelitian ini. WFCC merupakan pengembangan dari metode *Mel-Frequency Cepstral Coefficients* (MFCC) dimana pada tahapannya, *Fast Fourier Transform* (FFT) digantikan menggunakan *Wavelet Transform*. Tahapan ekstraksi fitur WFCC terdapat pada diagram berikut.



Gambar 1. Diagram Alir WFCC

A. Pre-emphasis

Tahap *pre-emphasis* merupakan tahap pertama yang berfungsi untuk filterisasi sinyal suara yang didapat dari ucapan manusia. Proses filterisasi ini dilakukan dengan cara mengeliminasi frekuensi sinyal yang diterima sehingga hanya tersisa sinyal dengan frekuensi tinggi saja, dimana dalam hal ini hanya sinyal suara ucapannya saja. Proses eliminasi ini dilakukan dengan tujuan untuk mengurangi *noise* yang dihasilkan dari suara yang masuk.

B. Frame Blocking

Tahap *frame blocking* atau yang biasa

disebut dengan *framing* merupakan tahapan dimana sinyal dibagi menjadi beberapa blok atau *frame* yang memiliki durasi lebih singkat, dikarenakan sinyal pada ucapan yang diterima selalu berubah. Panjang *frame* yang digunakan adalah antara 10 sampai 30 ms (Putra, 2017). Besaran ukuran dari *frame* diusahakan sepanjang mungkin agar dapat menampilkan resolusi frekuensi yang baik. Tetapi ukuran *frame* diusahakan sependek mungkin untuk dapat menampilkan resolusi waktu yang baik. Biasanya, proses *frame blocking* akan berlangsung dalam interval waktu 20 sampai 40 ms.

C. *Windowing*

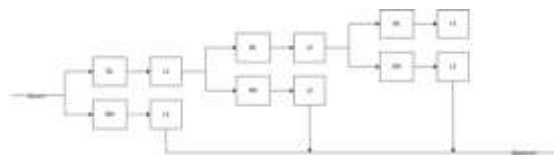
Pada tahap *windowing* dilakukan fungsi *weighting* pada setiap *frame* yang telah dibentuk pada tahap sebelumnya dengan tujuan untuk meminimalkan *discontinuities* pada ujung awal dan ujung akhir setiap *frame* (Rahmawanthi et al., 2019). Selain itu, *windowing* juga dilakukan untuk meminimalisir jumlah *sampling rate* yang rendah agar kebocoran spektral (*aliasing*) tidak terjadi (Putra, 2017).

Untuk mengurangi kebocoran spektral atau *aliasing*, hasil dari proses *framing* harus melewati fungsi *windowing* yang dirumuskan dalam persamaan (1), dimana N merupakan durasi dari *frame* sinyal suara yang didapatkan.

$$w(n) = 0,54 - 0,64 \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N - 1 \quad (1)$$

D. *Wavelet Transform*

Wavelet Transform merupakan proses pengubahan data dalam bentuk lain agar lebih mudah dianalisis (Sutarno, 2010). Pada *wavelet transform*, dilakukan pengubahan sinyal dalam berbagai *wavelet* basis dengan berbagai pergeseran dan penyekalaan, yang menyebabkan koefisien *wavelet* dari beberapa skala dapat dihitung dari koefisien *wavelet* pada resolusi tinggi berikutnya. *Wavelet Transform* memiliki beberapa jenis, Pada penelitian ini akan menggunakan *Discrete Wavelet Transform* (DWT). Pada proses DWT, sinyal dibagi menjadi dua bagian, frekuensi tinggi dan frekuensi rendah berturut- turut disertai *low-pass filter* dan *high-pass filter*. Frekuensi rendah dibagi kembali menjadi frekuensi tinggi dan frekuensi rendah, dan prosesnya terus berulang hingga sinyal tidak dapat didekomposisi lagi.



Gambar 2. Transformasi Wavelet Diskrit

Pada tahap Gabungan dari berbagai keluaran *high-pass filter* dan satu keluaran *low-pass filter* yang terakhir disebut sebagai koefisien *wavelet*, yang berisi informasi sinyal hasil transformasi yang telah terkompresi. Hasil dari proses transformasi menggunakan DWT kemudian akan disebut sebagai *spectrum*.

E. *Wavelet Frequency Warping*

Wavelet Frequency Warping dilakukan dengan menggunakan *Filterbank*. *Filterbank* sendiri merupakan salah satu format dari proses filter yang digunakan untuk mengetahui besaran energi dari beberapa *band frequency* tertentu pada sinyal suara. *Filterbank* pada penelitian ini diterapkan dalam domain frekuensi untuk keperluan WFCC.

F. *Discrete Concine Transform* (DCT)

Tahap ini merupakan tahap terakhir dari proses ekstraksi fitur pada WFCC, dimana sinyal yang dihasilkan dari proses sebelumnya akan dikonversi dari domain frekuensi menjadi domain waktu. Proses ini dilakukan untuk mendapatkan nilai koefisien dari hasil perkalian sinyal pada tahap sebelumnya yang sudah dikonversi ke domain waktu, yang selanjutnya akan menghasilkan log perkalian DCT yang nantinya akan menghasilkan *Wavelet-frequency Cepstrum Coefficient* (WFCC).

2.3 *K-Nearest Neighbor* (KNN)

Algoritma *K-Nearest Neighbor* merupakan algoritma yang digunakan untuk melakukan klasifikasi terhadap suatu objek yang dilakukan dengan cara mencari kelompok k buah data latih (*data training*) yang jaraknya paling dekat dengan objek pada data baru (*data testing*).

Menurut Adhi Putra (2021), Algoritma KNN memiliki prinsip yang sederhana, yakni bekerja berdasarkan jarak terpendek dari data uji ke data latih. Tahapan awal yang dilakukan pada algoritma KNN adalah menentukan parameter k , yakni jumlah *Neighbor* (tetangga) dengan jarak paling dekat.

Dalam algoritma KNN, pada umumnya jarak *Euclidean* antara dua objek dirumuskan

sebagai berikut, Dimana x dan y merupakan objek yang akan diukur jarak *Euclidean*-nya (Mustakim dan Giantika, F., 2016).

$$d_{xy} = \sqrt{\sum_{i=1}^N (x_i - y_i)^2} \quad (2)$$

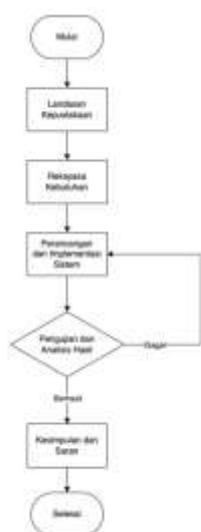
3. METODOLOGI

3.1 Tipe Penelitian

Penelitian ini termasuk kedalam tipe penelitian implementatif pengembangan. Implementatif berarti mewujudkan suatu metode atau teori menjadi sebuah bentuk yang nyata yang dapat bekerja sebagaimana mestinya. Pengembangan adalah proses untuk meningkatkan suatu hal agar menghasilkan keluaran yang lebih baik dibandingkan dengan keluaran terdahulu.

3.2 Metodologi Penelitian

Tahapan Metodologi Penelitian ini dapat dilihat pada Gambar 2 berikut.



Gambar 3. Diagram Alir Metodologi Penelitian

Tahapan pertama yang dilakukan adalah landasan kepustakaan, dimana peneliti melakukan studi literatur untuk mengumpulkan informasi dari penelitian sebelumnya sebagai dasar teori pada penelitian ini.

Pada tahap selanjutnya, peneliti melakukan proses rekayasa kebutuhan terkait apa saja yang diperlukan dalam penelitian ini yang terbagi menjadi dua, yakni kebutuhan perangkat keras dan kebutuhan perangkat lunak. Untuk kebutuhan perangkat lunak, yang dibutuhkan meliputi sistem operasi dan bahasa pemrograman. Untuk kebutuhan perangkat

keras, yang dibutuhkan meliputi mikrokomputer, LCD, dan *microphone*.

Pada tahap selanjutnya peneliti melakukan proses perancangan dan implementasi sistem. Proses pertama yang dilakukan adalah ekstraksi fitur dari dataset yang sudah ada dengan menggunakan WFCC yang hasilnya akan berbentuk file dengan format *xlsx*.

Dari dataset yang telah terekstraksi, selanjutnya akan dibuat menjadi *data training* yang akan menghasilkan model emosi marah untuk setiap intensitas dan kemudian akan dimuat ke dalam Raspberry Pi 4. Selanjutnya, sampel suara akan diambil melalui *microphone* untuk dibandingkan dengan model yang sudah dibuat sebelumnya dan akan diklasifikasikan menggunakan metode KNN.

Tahapan selanjutnya merupakan proses pengujian untuk mengetahui tingkat akurasi sistem yang telah dibuat dan untuk menguji apakah sistem yang dibuat sudah sesuai dengan tujuan penelitian. Peneliti melakukan pencocokan dari setiap model yang telah disimpan dengan suara yang diterima oleh *microphone*.

Pada tahap terakhir, peneliti menarik kesimpulan dari penelitian ini dengan memaparkan hasil uji yang telah dilakukan serta seberapa besar tingkat akurasi yang dihasilkan. Selanjutnya, peneliti akan memberikan penjelasan mengenai apa saja yang dapat dikembangkan dari penelitian ini pada bagian saran, yang bertujuan untuk membantu pengembangan topik ini agar berguna di masa depan.

4. ANALISIS KEBUTUHAN

4.1 Kebutuhan Fungsional

1. Sistem dapat menerima suara melalui *microphone* yang terhubung dengan Raspberry Pi 4.
2. Sistem dapat melakukan ekstraksi, dan klasifikasi dari suara yang diterima.
3. Raspberry Pi 4 dapat terhubung dengan LCD dan menampilkan hasil klasifikasi.

4.2 Spesifikasi Sistem

1. Sistem membutuhkan perangkat yang seluruhnya bersifat *portable* dan dapat terkoneksi satu sama lain melalui kabel maupun *wireless*.
2. Sistem harus memiliki memori yang cukup untuk menyimpan file suara

sementara yang telah diterima.

4.3 Analisis Kebutuhan Perangkat Keras dan Perangkat Lunak

Kebutuhan Perangkat Keras dan Perangkat Lunak yang didasarkan pada kebutuhan fungsional dan spesifikasi sistem dapat dilihat pada Tabel 1 berikut.

Tabel 1. Analisis Kebutuhan Perangkat

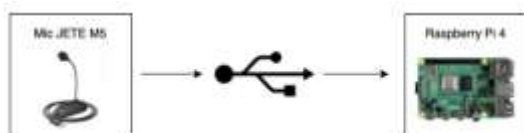
Jenis	Kebutuhan	Perangkat
Perangkat Keras	Perangkat mikrokomputer <i>portable</i>	Raspberry Pi 4 Model B
	LCD yang dapat terhubung dengan Mikrokomputer	LCD HDMI 7" inch
	<i>Microphone</i>	JETE M5
Perangkat Lunak	Kode proses ekstraksi dan klasifikasi	<i>Library Scikit-learn Library PyWavelets</i>
	Perangkat Lunak perancang aplikasi <i>smartphone</i>	MIT App Inventor

5. HASIL DAN PEMBAHASAN

5.1 Perancangan

Perancangan sistem pada penelitian ini terdiri dari rancangan perangkat keras dan rancangan perangkat lunak. Proses perancangan dibuat untuk mempermudah tahap implementasi sistem agar lebih terarah dan terstruktur.

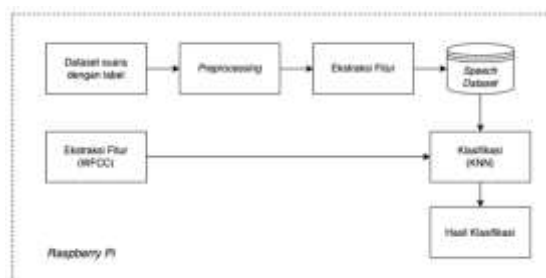
Perangkat Perekam Suara dirancang dengan menggunakan *microphone* JETE M5 dan mikrokomputer Raspberry Pi 4 Model B. Suara yang telah diterima dari *microphone* akan direkam dan disimpan di dalam memori mikrokomputer dalam bentuk format .wav. Untuk menghubungkan *microphone* JETE M5 dengan mikrokomputer Raspberry Pi 4, akan digunakan kabel dengan port USB yang dijelaskan dalam Gambar 4 berikut.



Gambar 4. Diagram Skematik Perancangan Perangkat Keras Perekaman Suara

Proses ekstraksi dan klasifikasi suara dirancang untuk dilakukan seluruhnya di dalam mikrokomputer Raspberry Pi 4. Pada tahapan

awal, diperlukan model yang diambil dari dataset yang sudah tersedia. Dataset tersebut kemudian diproses dengan cara diekstraksi menggunakan WFCC dan kemudian diklasifikasikan menggunakan algoritma KNN, hasil klasifikasi tersebut akan menjadi model yang akan dimasukkan ke dalam mikrokomputer untuk selanjutnya digunakan sebagai tolak ukur kelas intensitas emosi marah yang akan dideteksi. Kelas yang diberikan adalah HIGH (*Rage*), MIDDLE (*Anger*), dan LOW (*Annoyance*).



Gambar 5. Diagram Blok Perancangan Ekstraksi dan Klasifikasi Suara

5.2 Implementasi

Proses implementasi sistem berupa pengimplementasian perangkat keras dan perancangan perangkat lunak dari sistem pengenalan intensitas emosi marah melalui ucapan menggunakan ekstraksi WFCC dan algoritma KNN pada Raspberry Pi 4 Model B.

Sesuai dengan perancangan perangkat keras perekaman suara dimana *microphone* akan dihubungkan dengan mikrokomputer Raspberry Pi 4 melalui port USB, maka pengimplementasiannya ditampilkan pada gambar 8 berikut.



Gambar 6. Implementasi Perangkat Keras Perekaman Suara

Pada tahapan awal implementasi ekstraksi dan klasifikasi suara, peneliti membuat kode program untuk membuat model yang akan dipakai dalam proses klasifikasi sistem.

Dalam pengimplementasian Transformasi Wavelet pada ekstraksi WFCC, digunakan library PyWavelets melalui fungsi `pywt.wavdec()` untuk memproses data suara yang diberikan dari tahapan tahapan pada fungsi sebelumnya.

5.3 Pengujian

Pada pengujian ini, akan dihasilkan informasi mengenai ketepatan hasil akurasi ekstraksi WFCC dan klasifikasi KNN pada sistem deteksi intensitas emosi marah. Untuk memperoleh hasil tersebut, dilakukan perhitungan menggunakan persamaan 3 berikut.

$$akurasi (\%) = \frac{jumlah\ data\ benar}{total\ data} \times 100\% \quad (3)$$

Setelah dilakukan pengujian sistem, didapatkan hasil sebagai berikut.

Tabel 2 Hasil Uji Sampel Suara pada Sistem

Sampel	Label	Hasil Deteksi	Keterangan
1051_ANG_HI	High	High	Berhasil
1041_ANG_HI	High	High	Berhasil
1054_ANG_HI	High	High	Berhasil
1076_ANG_HI	High	Mid	Gagal
1026_ANG_HI	High	High	Berhasil
1004_ANG_MD	Mid	Mid	Berhasil
1019_ANG_MD	Mid	Low	Gagal
1025_ANG_MD	Mid	Low	Gagal
1033_ANG_MD	Mid	Mid	Berhasil
1090_ANG_MD	Mid	Low	Gagal
1012_ANG_LO	Low	Low	Berhasil
1019_ANG_LO	Low	Low	Berhasil
1044_ANG_LO	Low	Mid	Gagal
1051_ANG_LO	Low	Low	Berhasil
1084_ANG_LO	Low	Low	Berhasil

Berdasarkan hasil pengujian yang ditampilkan pada Tabel 6.1, diketahui bahwa sistem berhasil memprediksi secara benar 10 dari 15 sampel data sehingga akurasi yang diperoleh adalah sebesar 66,67%.

Berdasarkan hasil pengujian, ditetapkan bahwa keseluruhan sistem berfungsi dengan baik mulai dari fungsional perangkat keras, hingga fungsional perangkat lunak mulai dari proses ekstraksi hingga menampilkan hasil deteksi pada

aplikasi. Hal ini menunjukkan bahwa spesifikasi sistem telah memenuhi kebutuhan fungsional sistem yang telah dibuat.

Hasil Pengujian pada akurasi dalam mendeteksi emosi marah melalui suara menunjukkan bahwa metode ekstraksi WFCC dan klasifikasi KNN cukup efektif dalam mendeteksi intensitas emosi marah yang tinggi atau *Rage* (HIGH). Hal ini terjadi karena jenis suara yang memiliki tingkat emosi HIGH cenderung memiliki frekuensi yang bervariasi terhadap waktu, yang menjadi keunggulan dari Transformasi Wavelet pada ekstraksi WFCC. Untuk pengujian intensitas emosi marah pada tingkatan sedang (MID) dan rendah (LOW) terdapat beberapa kegagalan yang kemungkinan terjadi karena kurangnya frekuensi yang bervariasi terhadap waktu, sehingga terlihat mirip dan bisa saling bertukar antara sedang dan rendah, yang juga dapat dilihat pada hasil bahwa pendeteksian yang gagal tidak pernah menghasilkan jenis emosi tinggi (HIGH).

6. KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan tahapan penelitian yang didasarkan pada hasil perancangan, implementasi, dan pengujian yang telah dilakukan sebelumnya, diperoleh kesimpulan sebagai berikut.

1. Metode Ekstraksi WFCC (*Wavelet-Based Frequency Cepstral Coefficients*) dan Klasifikasi Algoritma KNN (*K-Nearest Neighbor*) cukup efektif dalam mendeteksi intensitas emosi marah *rage* atau tinggi (HIGH) dengan akurasi sebesar 66,67%

6.2 Saran

Berdasarkan hasil dan kesimpulan yang didapat, saran dari penelitian ini adalah sebagai berikut.

1. Menambahkan filter yang dapat memberikan lebih banyak variasi frekuensi terhadap waktu, sehingga hasil prediksi akurasi suara akan menjadi lebih akurat.

7. DAFTAR PUSTAKA

- Abdalla, M.I. and Ali, H.S., 2010. Wavelet-Based Mel-Frequency Cepstral Coefficients for Speaker Identification using Hidden Markov Models. *JOURNAL OF TELECOMMUNICATIONS*, 1(2),

- pp.16–21.
- Adam, T.B., Salam, M.S. and Gunawan, T.S., 2013. Wavelet based Cepstral coefficients for neural network speech recognition. *2013 IEEE International Conference on Signal and Image Processing Applications*.
- Adhi Putra, A.D., 2021. Analisis Sentimen Pada Ulasan Pengguna aplikasi bibit Dan Bareksa Dengan Algoritma knn. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 8(2), pp.636–646.
- Anant, K.S. and Dowla, F.U., 1997. Wavelet transform methods for phase identification in three-component seismograms. *Bulletin of the Seismological Society of America*, 87(6), pp.1598–1612.
- Anastasya, N., 2020. *Pengelompokan Musik Berdasarkan Emosi Menggunakan Metode Transformasi Haar Wavelet*. thesis. Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, pp.10524–10530.
- Davis, C.P., 2021. *Medical definition of anger*. [online] MedicineNet. Available at: <<https://www.medicinenet.com/anger/definition.htm>> [Accessed 30 Aug. 2022].
- Gumelar, A.B., Yuniarno, E.M., Anggraeni, W., Sugiarto, I., Kristanto, A.A. and Purnomo, M.H., 2020. Kombinasi Fitur Multispektrum Hilbert dan Cochleagram untuk Identifikasi Emosi Wicara. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 9, pp.180–189.
- Helmiyah, S., Fadlil, A. and Yudhana, A., 2019. Pengenalan Pola Emosi Manusia Berdasarkan UCAPAN Menggunakan Ekstraksi fitur Mel-frequency cepstral coefficients (MFCC). *CogITo Smart Journal*, 4(2), p.372.
- Lei, H., 2022. [online] Joint Factor Analysis (JFA) and I-vector tutorial - ICSI. Available at: <https://www1.icsi.berkeley.edu/Speech/presentations/AFRL_ICSI_visit2_JFA_tutorial_icsitalk.pdf> [Accessed 26 Aug. 2022].
- Mustakim and F, G.O., 2016. Algoritma K-Nearest Neighbor Classification Sebagai Sistem Prediksi Predikat Prestasi Mahasiswa. *Jurnal Sains, Teknologi dan Industri*, 13(2), pp.195–202.
- Nasersharif, B. and Akbari, A., 2014. Application of wavelet transform and wavelet thresholding in robust sub-band speech recognition. pp.345–348.
- Putra, D. and Resmawan, A., 2011. Verifikasi biometrika suara menggunakan metode MFCC dan DTW. *Lontar Komputer*, 2(1), pp.8–21.
- Rahmawanthi, I., Raharjo, J. and Rusdinar, A., 2019. Deteksi Suara Manusia Dalam Keadaan Emosi Dengan Menggunakan Linear Predictive Coding (LPC) Dengan Klasifikasi Coarse To Fine Search (CFS) Berbasis Pengolahan Data. *e-Proceeding of Engineering*, 6, pp.656–663.
- Sulistyaningrun, D.R. and Khukmiati, H., 2004. Penerapan Transformasi Wavelet diskrit untuk reduksi noise pada citra digital. *Limits: Journal of Mathematics and Its Applications*, 1(1), pp.49–57.
- Sundawa, A.A., Putrada, A.G. and Suwastika, N.A., 2019. Implementasi dan Analisis Simulasi Deteksi Emosi Melalui Pengenalan Suara Menggunakan Mel-Frequency Cepstrum Coefficient dan Hidden Markov Model Berbasis IOT. *e-Proceeding of Engineering*, 6, pp.2100–2107.
- Sutarno, 2010. Analisis Perbandingan Transformasi Wavelet pada Pengenalan Citra Wajah. *Jurnal Generic*, 5(2), pp.15–21.
- Tzanetakis, G., Essl, G. and Cook, P., 2001. Audio Analysis using the Discrete Wavelet Transform. *Proceedings of the Conference in Acoustics and Music Theory Applications*, pp.318–323.
- Vergin, R. and O'Shaughnessy, D., 1995. Pre-emphasis and speech recognition. *Proceedings 1995 Canadian Conference on Electrical and Computer Engineering*, 2, pp.1062–1065.
- Yahya, A., 2012. Deteksi Emosi Melalui Pengenalan Suara Menggunakan Linear Predictive Coding (LPC) dan Hidden Markov Model (HMM).