



Identifikasi Jenis Kelamin Secara Real Time Berdasarkan Suara Pada Raspberry Pi

Mirza Ardiana¹, Titon Dutono² dan Tri Budi Santoso³

¹Politeknik Elektronika Negeri Surabaya, Teknik Elektro, email: mirza@pasca.student.pens.ac.id

²Politeknik Elektronika Negeri Surabaya, Teknik Telekomunikasi, email: titon@pens.ac.id

³Politeknik Elektronika Negeri Surabaya, Teknik Telekomunikasi, email: tribudi@pens.ac.id

Abstrak

Suara setiap penutur memiliki karakter spesifik yang unik, dipengaruhi oleh jenis kelamin, usia, emosi, dialek, dll. Penggunaan identifikasi jenis kelamin berdasarkan suara berkembang pesat, seperti pada bidang sistem keamanan, pengenalan ucapan, kecerdasan buatan, dll. Tetapi dalam pengolahan suara terdapat kesulitan yang mana karakteristik sinyal suara berdasarkan bertambahnya usia tergolong sulit dalam penentuan akurasi, dan terdapat nilai frekuensi fundamental yang *overlapping* di antara laki-laki dan perempuan. Penelitian ini telah dilakukan suatu pemodelan sistem identifikasi jenis kelamin berdasarkan suara secara *real-time* pada perangkat Raspberry Pi. Sistem ini diimplementasikan dengan 2 metode yaitu algoritma YIN dan ekstraksi fitur ciri Mel-Frequency Cepstral Coefficient (MFCC). Hasil pengujian menunjukkan keberhasilan identifikasi pada tuning parameter skema kedua lebih baik daripada skema pertama dengan mempersempit parameter frekuensi *overlapping*. Sistem ini menggunakan data uji dan data latih sebanyak 50 wav suara perempuan dan 50 wav suara laki-laki. Pada data uji perempuan di closed test akurasinya dari 98% menjadi 100%, lalu pada open test bermula 92% menjadi 96%. Sedangkan pengujian data laki-laki closed test dari 92% naik menjadi 98%, dan open test semula 90% naik menjadi 94%. Hal ini menunjukkan bahwa pada data yang digunakan pada penelitian ini lebih cocok menggunakan *tuning* parameter skema kedua untuk meningkatkan hasil akurasi.

Kata kunci: *real-time, gender, YIN, MFCC, Raspberry Pi*

Abstract

The voice of each speaker has a unique specific character, influenced by gender, age, emotion, dialect, etc. The use of voice-based gender identification is growing rapidly, such as in the fields of security systems, speech recognition, AI, etc. However, there are difficulties in sound processing because with increasing age, it will affect the character of the sound signal, and there is also an overlapping frequency between male and female. This research, real-time gender identification system using voice on Raspberry Pi has been carried out. System implemented by 2 methods, namely YIN algorithm and MFCC. The test results show that identification success of the second scheme parameter tuning is better than the first scheme by narrowing the overlapping frequency parameters. This system uses test data and train data of 50 female and 50 male voices. In the female test data in the closed test, the accuracy is from 98% to 100%, then open test starts from 92% to 96%. Meanwhile, the male closed test data

increased from 92% to 98%, and open test was originally 90% increased to 94%. This shows that data used in this study is more suitable to use the second scheme parameter tuning to improve accuracy results.

Keywords: *real-time, gender, voice, Raspberry Pi*

1. Pendahuluan

Suara adalah sebuah data yang mengandung informasi, salah satunya adalah informasi mengenai jenis kelamin [1]. Pada dasarnya setiap individu memiliki karakteristik suara yang spesifik, sehingga bisa membedakan dengan melakukan analisa suara dari satu individu dengan individu lainnya. Identifikasi jenis kelamin berdasarkan suara adalah mengidentifikasi jenis kelamin dari analisa suara penuturnya. Dewasa ini perkembangan identifikasi jenis kelamin berdasarkan suara sangat penting dalam beberapa bidang, diantaranya dalam bidang *speaker identification, speech recognition, emotion recognition*, dll [2].

Suara memiliki keunikan tersendiri dengan mempunyai sebuah parameter nilai level dengan rentang tertentu dan nilai frekuensi yang bervariasi, adapun suara juga dapat dijangkau dan tidak dapat dijangkau oleh pendengaran manusia [3]. Berdasarkan teori yang dikemukakan Furui bahwa "Frekuensi Fundamental suara manusia laki-laki dan perempuan adalah dua buah kurva terdistribusi normal dengan sedikit *overlapping*, yang berasal dari nilai rata-rata dan simpangan baku, dari banyak populasi berbeda" [4]. Hal ini yang membuat tantangan dalam melakukan prediksi jenis kelamin berdasarkan suara dengan menggunakan parameter pitch, karena terdapat nilai frekuensi *overlapping* diantara laki-laki dan perempuan yang membutuhkan analisa lebih mendalam untuk memprediksinya. Karakteristik dari frekuensi fundamental (*pitch*) pada suara manusia secara umum berkurang seiring dengan pertambahan usia.

Membanjirnya perangkat mini PC yang terjangkau dari sisi harga dan kemudahan di dalam memperolehnya memberikan kesempatan kepada peneliti untuk mengimplementasikan berbagai model dalam uji *real-time*. Salah satu bentuknya implementasi platform Raspberry Pi yang sudah banyak digunakan yaitu dalam berbagai perangkat teknologi. Raspberry Pi merupakan sebuah perangkat dari komputer berbentuk kecil dengan kemampuan penggunaan di berbagai bidang. Platform ini diharapkan dapat mendukung dalam implementasi identifikasi jenis kelamin berdasarkan suara secara *real-time* dengan menggunakan perangkat *microphone* sebagai media input suara dan perangkat *soundcard adapter* sebagai media perekaman dan pemutaran suara.

Berdasarkan penelitian sebelumnya, proses identifikasi jenis kelamin telah menggunakan berbagai metodologi ataupun berbagai aplikasi. Dalam melakukan prediksi jenis kelamin dengan menggunakan suara dengan karakteristik perbedaan fisiologis (glottis, ketebalan saluran vokal) dan perbedaan dalam fonetik [5,6,7,8,9,10]. Adapun fitur yang sering digunakan dalam ekstraksi suara adalah dengan menggunakan fitur *Pitch* [7] dan Forman [8]. Estimasi nilai Pitch dari suara dengan menggunakan algoritma YIN berdasarkan metode autokolerasi dengan beberapa modifikasi yang terjadi untuk mengurangi kesalahan prediksi [11]. Melakukan analisa suara dengan menggunakan fitur ekstraksi ciri *Mel-Frequency Cepstral Coefficient* (MFCC) [12,13,14]. Pada dasarnya MFCC merupakan sebuah ekstraksi fitur dengan menggunakan *coefficient Mel-Frequency Cepstral* yang dimodelkan berdasarkan persepsi suara telinga manusia dan bukan dalam bentuk skala linier [13]. Adapun klasifikasi dengan prinsip perhitungan jarak dalam data MFCC lebih baik menggunakan Mahalanobis Distance dibandingkan Euclidean Distance, karena dengan Mahalanobis Distance bisa memberikan kenaikan dari performa klasifikasi dan mengurangi kesalahan klasifikasi [15,16].

Maka penelitian ini mengusulkan suatu sistem *real-time* proses identifikasi jenis kelamin berdasarkan suara penutur dengan menggunakan media Raspberry Pi. Metode yang digunakan melalui 2 macam yaitu dengan menggunakan fitur pitch dan fitur koefisien MFCC. Keutamaan dari penelitian ini adalah berupa perwujudan sistem secara *real-time* dengan menggunakan 2 tahapan metode serta melakukan *tuning* data untuk menghasilkan identifikasi jenis kelamin berdasarkan suara yang benar dan mengurangi kesalahan.

2. Metodologi

Sistem identifikasi jenis kelamin berdasarkan suara dari penutur dengan menggunakan media Raspberry Pi melalui beberapa tahapan proses diantaranya melakukan persiapan alat, kemudian berlanjut pada pengumpulan data, proses pembelajaran data, dan dilanjutkan dengan proses pengujian data.

2.1 Mempersiapkan alat penelitian

Pada penelitian ini membutuhkan beberapa alat yang digunakan. Media utama dalam penelitian ini menggunakan Raspberry Pi 3 Model B V1.2 sebagai komputer *single-board* yang berukuran kecil yang bisa digunakan untuk menjalankan program tertentu dengan bahasa python. Media Raspberry Pi dipasangkan dengan *Soundcard* UGREEN USB External, di mana perangkat ini berfungsi untuk media perekaman dan pemutaran suara. *Soundcard* dihubungkan dengan *earphone* atau *headphone* pada jalur input *soundcard* untuk masukan suara penutur. Keluaran dari sistem ini bisa ditampilkan dengan menggunakan layar monitor dan *speaker* eksternal.

2.2 Proses Pengumpulan Data

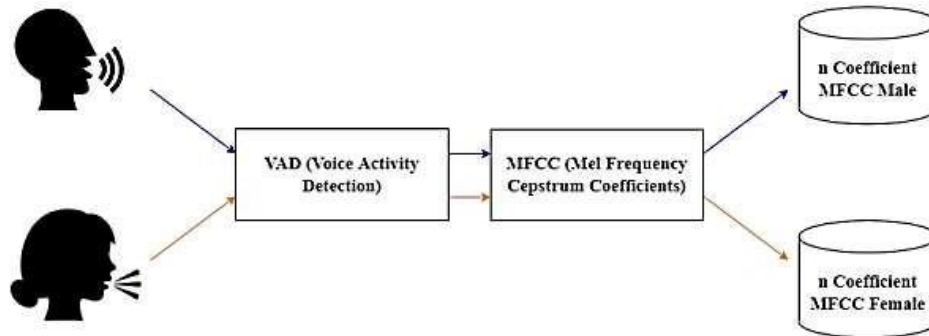
Penelitian ini menggunakan data suara dari penutur perempuan dan suara penutur laki-laki dengan proses perekaman suara dalam keadaan bebas *noise*. Koresponden penutur dibatasi pada penutur yang berusia 19 – 25 tahun. Hal ini mengacu pada penelitian yang telah dilakukan sebelumnya, yang mana telah didapatkan data bahwa pada rentang usia 19-25 tahun suara manusia dapat menghasilkan frekuensi pitch yang konstan, jika dibandingkan dengan suara manusia pada usia diluar rentang tersebut [12]. Proses perekaman data dilakukan pada media Raspberry Pi dengan kondisi sudah terpasang *microphone* beserta *soundcard*. Perekaman suara dilakukan secara *real-time* dengan menerapkan prinsip *Voice Activity Detection* (VAD). Pada dasarnya suatu sinyal suara terdapat informasi yang terdiri dari periode aktif (*voice*) dan periode pasif (*silent* atau *noise*), untuk itu pemanfaatan metode VAD digunakan untuk identifikasi status aktivitas pada sinyal suara, apakah berupa periode aktif atau periode pasif pada setiap frame. Hal terpenting dalam VAD adalah ketepatan dalam pembuatan keputusan terhadap setiap frame dari sinyal speech apakah frame tersebut berupa frame aktif atau frame pasif (*silent* dan/atau *background noise*). Adapun proses perekaman suara penutur seringkali terdapat noise akibat kondisi tempat saat perekaman atau berasal dari kondisi perangkat yang digunakan. Hal ini sudah diantisipasi dengan penggunaan algoritma YIN pada sistem ini, yang akan memproses suara dengan menggunakan prinsip *threshold* bawah dan atas pada frekuensi fundamental, sehingga selain batasan *threshold* tersebut akan dinormalisasi menjadi nol.

2.3 Proses Pembelajaran Data

Proses *training* atau yang disebut pembelajaran merupakan sebuah tahapan untuk mendapatkan nilai koefisien (*Mel-Frequency Cepstral Coefficient*) MFCC, dimana pada penelitian kali ini menggunakan 12 koefisien MFCC dari setiap wav suara yang di proses dan disimpan dalam bentuk file berekstensi csv. Hasil data ini dijadikan *database* untuk menentukan nilai jarak terdekat dari data test terhadap *database* yang nanti akan mempengaruhi hasil klasifikasinya. Adapun pencarian koefisien MFCC pada penelitian ini berjumlah 13 koefisien, akan tetapi koefisien yang diproses hanya koefisien ke satu sampai terakhir saja (12 koefisien). Hal ini

berlandaskan pada penelitian yang pernah dilakukan bahwa koefisien ke nol mengindikasikan energi dari frame sinyal sehingga tidak *reliable*. Secara sederhana gambaran proses pembelajaran dapat disajikan pada Gambar 1.

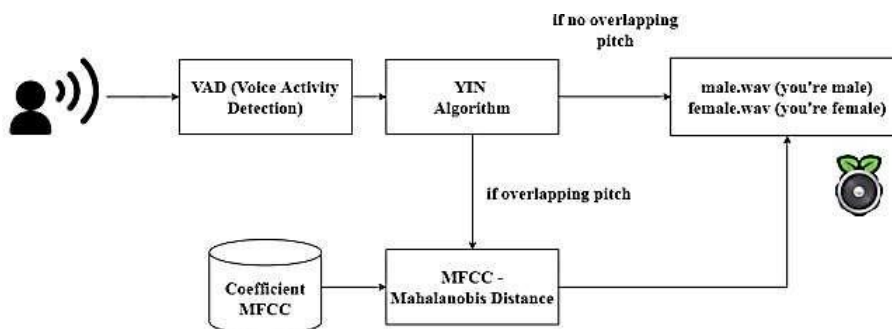
Pada tahapan proses pembelajaran menggunakan 10 penutur perempuan dan 10 penutur laki-laki. Setiap penutur akan mengucapkan 5 kalimat yang sudah ditentukan, sehingga jumlah data training yang digunakan untuk menghitung jarak data berjumlah 50 data wav MFCC perempuan dan 50 data wav MFCC laki-laki. Data pembelajaran ini berfungsi untuk melakukan klasifikasi data test baru dengan menggunakan prinsip *Mahalanobis Distance*.



Gambar 1. Alur sistem proses pembelajaran data

2.4 Proses Pengujian Data

Proses pengujian berawal dari suara penutur yang dilakukan proses perekaman pada media Raspberry Pi dengan menggunakan prinsip VAD (*Voice Activity Detection*). Kemudian sistem akan melakukan pemrosesan dengan mendapatkan informasi nilai pitch atau frekuensi fundamental dari suara data tes dengan menggunakan algoritma YIN. Secara sederhana dapat dilihat seperti pada Gambar 2.



Gambar 2. Alur sistem proses pengujian data

Pada dasarnya algoritma YIN melakukan estimasi nilai frekuensi fundamental dari suara wicara, adapun proses dalam algoritma YIN terdiri dari fungsi autokolerasi, fungsi difference, proses CMNDF (*Cumulative Mean Normalized Difference Function*), kemudian dilanjutkan dengan proses pemberian absolute threshold, interpolasi, dan melakukan estimasi nilai pitch atau frekuensi fundamental [11]. Apabila terdapat data test baru, maka sistem akan melakukan perhitungan nilai pitch dari suara dan menggolongkan ke 3 pilihan kategori yaitu apakah termasuk dalam pitch laki-laki, pitch perempuan, atau pitch kategori tengah atau disebut *overlapping*. Ketika didapatkan nilai pitch *overlapping*, maka sistem akan melanjutkan ke proses selanjutnya yaitu mencari nilai koefisien ekstraksi ciri MFCC. Jika tidak maka sistem akan langsung memberikan identifikasi jenis kelamin berdasarkan kategori nilai pitch.

Pada data test yang tergolong pitch kategori tengah atau *overlapping*, sistem akan melanjutkan proses untuk mendapatkan nilai ekstraksi ciri MFCC dari data test. Pada dasarnya MFCC merupakan salah satu fitur ekstraksi ciri dengan menggunakan koefisien *Mel-Frequency Cepstral*, dimana proses dalam MFCC terdiri dari melakukan konversi suara dari analog ke digital terlebih dahulu, kemudian dilanjutkan dengan proses pre-emphasize filtering, frame blocking, windowing, analisa fourier, DFT, FFT *Fast Fourier Transform* (FFT), Mel Frequency Warping, *Discrete Cosine Transform* (DCT), dan proses cepstral filtering. Dari tahapan ini sistem akan mendapatkan koefisien MFCC untuk dilakukan perhitungan jarak terhadap *database* MFCC laki-laki dan perempuan dengan menggunakan Mahalanobis Distance.

Adapun Mahalanobis Distance merupakan proses menghitung jarak dua titik yang melibatkan kovarians atau kolerasi antar peubah yang dinyatakan dalam bentuk matrik dan vector. Berikut ini formula dalam menghitung Mahalanobis Distance:

$$d_{ij}^2 = (x_i - x_j)^T \Sigma^{-1} (x_i - x_j) \quad (1)$$

Prinsip perhitungan klasifikasi dengan perhitungan Mahalanobis Distance yaitu semakin dekat jarak data test terhadap salah satu *database* MFCC, maka sistem akan melakukan klasifikasi berdasarkan nilai jarak terdekat. Output dari sistem ini selain hasil klasifikasi jenis kelamin yang ditampilkan pada monitor, juga dilengkapi dengan pemutaran wav yang bersuarakan “*You are voice is male !!*” pada file bernama male.wav dan bersuarakan “*You are voice is female !!*” pada file female.wav sesuai dengan klasifikasi yang dihasilkan oleh sistem.

3. Hasil dan Pembahasan

Proses pengujian dalam mengidentifikasi jenis kelamin secara *real-time* pada media Raspberry Pi ini dilakukan dengan menggunakan data test berjumlah 10 penutur perempuan dan 10 penutur laki-laki. Setiap penutur mengucapkan 5 kalimat yang berbeda untuk setiap percobaan, sehingga pada penelitian ini menggunakan 50 data wav suara perempuan, dan 50 data wav suara laki-laki sebagai data pengujian. Skema pengujian ini terbagi menjadi dua macam, yaitu untuk skema pertama menggunakan parameter frekuensi *overlapping* pada rentang 150 – 210 Hz. Sedangkan pada skema yang kedua pengujian dengan menggunakan parameter frekuensi *overlapping* yang dipersempit menjadi rentang 180 – 210 Hz.

3.1 Skema Pengujian Pertama

Pada skema pertama menggunakan rentang frekuensi *overlapping* 150 - 210 Hz. Pada pengujian ini dibedakan menjadi dua tipe yaitu *closed test* dan *open test*. *Closed test* merupakan data set pada proses pembelajaran MFCC yang telah digunakan sebelumnya digunakan lagi menjadi data test. Sedangkan *open test* menggunakan data test yang baru untuk dilakukan pengujian. Sehingga pada skema pertama ini terdapat 50 data wav suara perempuan dan 50 data wav laki-laki di setiap pengujian *closed test* maupun *open test*. Maka berikut ini hasil pengujian yang telah dilakukan secara *real-time*, yaitu sebagai berikut:

Tabel 1. Hasil uji data perempuan skema 1

Jenis Pengujian	Identifikasi Benar	Identifikasi Salah	Akurasi
<i>Closed test</i>	49	1	98%
<i>Open test</i>	46	4	92%

Tabel 2. Hasil uji data laki-laki skema 1

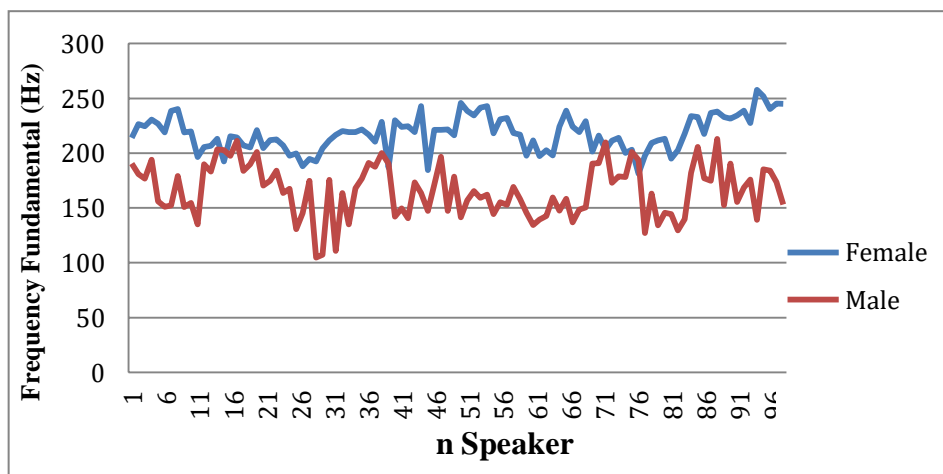
Jenis Pengujian	Identifikasi Benar	Identifikasi Salah	Akurasi
<i>Closed test</i>	46	4	92%
<i>Open test</i>	45	5	90%

Berdasarkan tabel pengujian, baik pada penutur laki-laki dan penutur perempuan diperoleh hasil akurasi yang cukup bagus. Pada penutur perempuan dengan jenis pengujian *closed test* mendapatkan akurasi sebesar 98% dan pada *open test* mendapatkan akurasi sebesar 92%. Sedangkan pada pengujian penutur laki-laki dengan jenis pengujian *closed test* mendapatkan akurasi sebesar 94% dan pada *open test* mendapatkan akurasi senilai 90%. Adapun hasil identifikasi salah pada data uji skema pertama ini dikarenakan beberapa data tersebut tergolong dalam kategori suara lain maupun termasuk overlapping yang sulit untuk diidentifikasi. Dimana masih terdapat beberapa wav suara yang tergolong overlapping yang mengalami kesalahan klasifikasi pada proses Mahalanobis Distance, sehingga data uji mengalami error atau kesalahan klasifikasi.

Berdasarkan metode sistem yang digunakan, dengan algoritma YIN didapatkan nilai pitch atau nilai frekuensi fundamental dari suara penutur, apabila didapatkan nilai pitch kategori tengah atau disebut dengan *overlapping* maka sistem akan melanjutkan proses untuk mencari nilai fitur MFCC yang kemudian dilakukan penghitungan jarak data dengan data set menggunakan Mahalanobis Distance. Pada pengujian pertama ini menggunakan 3 keputusan berdasarkan nilai pitch pada proses Algoritma Yin, yaitu sebagai berikut :

- Jenis kelamin Laki-laki = 65 – 150 Hz
- Frekuensi Overlapping = 150 – 210 Hz
- Jenis kelamin Perempuan = 210 – 500 Hz

Sebaran data nilai pitch yang didapatkan pada 100 data test penutur perempuan dan 100 penutur laki-laki dengan menggunakan algoritma YIN, seperti disajikan pada gambar berikut:



Gambar 3. Persebaran Data Nilai Pitch

Dari gambar tersebut terlihat jelas bahwa diantara pitch penutur laki-laki dan penutur perempuan terdapat rentang nilai yang saling *overlapping*, hal ini sesuai dengan teori Furui yang disampaikan [4]. Berdasarkan hasil pengujian data test yang termasuk dalam kategori pitch *overlapping*, maka sistem melakukan pencarian nilai fitur ekstraksi ciri MFCC dengan klasifikasi menggunakan perhitungan jarak *Mahalanobis Distance*, yang merupakan cara penentuan klasifikasi dengan menggunakan covarian matrix dalam perhitungan jarak nya sehingga lebih kompleks daripada penggunaan fungsi *Euclidean Distance*.

3.2 Skema Pengujian Kedua

Berdasarkan skema pengujian pertama, hasil identifikasi yang dihasilkan cukup bagus akan tetapi masih terdapat kesalahan yang dihasilkan. Pada pengujian skema kedua dilakukan perubahan pada parameter frekuensi *overlapping* yang digunakan yaitu dengan mempersempit nilai kategori tengah atau *overlapping* menjadi 180 – 210 Hz dengan data set dan data test yang sama pada pengujian skenario pertama. Hasil data pengujian pada suara penutur perempuan dan laki-laki, seperti pada Tabel berikut:

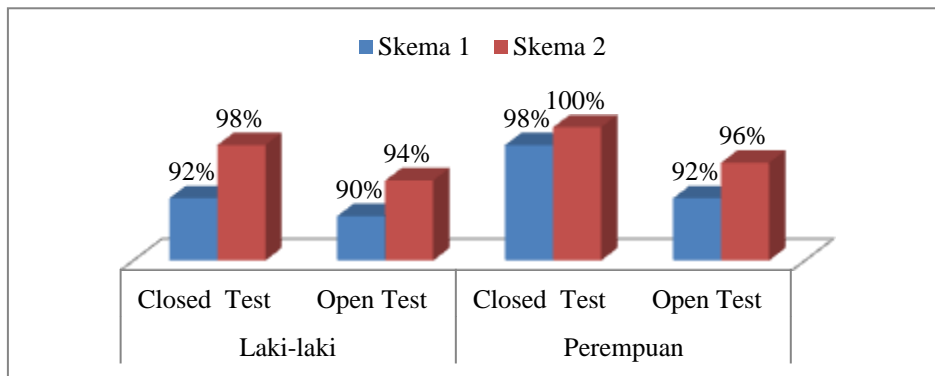
Tabel 3. Hasil uji data perempuan skema 2

Jenis Pengujian	Identifikasi Benar	Identifikasi Salah	Akurasi
<i>Closed test</i>	50	0	100%
<i>Open test</i>	48	2	96%

Tabel 4. Hasil uji data laki-laki skema 2

Jenis Pengujian	Identifikasi Benar	Identifikasi Salah	Akurasi
<i>Closed test</i>	49	1	98%
<i>Open test</i>	47	3	94%

Berdasarkan tabel di atas, dapat dijelaskan bahwa hasil pengujian data suara perempuan pada skema kedua ini menghasilkan akurasi *closed test* sebesar 100%, dan pada *open test* menghasilkan akurasi 96%. Sedangkan pada pengujian data suara laki-laki menghasilkan akurasi di *closed test* sebesar 98%, dan pada *open test* menghasilkan akurasi 94%. Adapun kesalahan klasifikasi pada pengujian skema kedua ini dikarenakan masih terdapat beberapa wav data uji yang termasuk kategori *overlapping* yang notabene nya sulit untuk diidentifikasi. Sehingga beberapa wav data uji ini masih belum berhasil diklasifikasikan oleh *Mahalanobis Distance*. Berdasarkan hasil data pengujian, pada skema kedua didapatkan akurasi yang meningkat daripada hasil pengujian skema pertama, dengan ilustrasi seperti berikut ini :



Te 4. Perbandingan hasil pengujian 2 skema

Pada data test perempuan di pengujian *closed test* yang semula 98% naik menjadi 100%, kemudian pada pengujian *open test* yang semula 92% naik menjadi 96%. Sedangkan pada pengujian data test laki-laki di pengujian *closed test* yang semula 92% naik menjadi 98%, kemudian pada pengujian *open test* yang semula 90% naik menjadi 94%. Berdasarkan penelitian ini, penggunaan *range* frekuensi pada algoritma YIN yang digunakan berpengaruh pada hasil akurasi yang dihasilkan. Dengan dipersempit nya nilai rentang *overlapping* pada data test yang digunakan, maka mempengaruhi data suara tersebut diputuskan untuk dilanjutkan ke tahapan selanjutnya atau tidak. Sehingga pada algoritma YIN, ketika sudah didapatkan nilai pitch diluar frekuensi *overlapping* maka bisa langsung terklasifikasi jenis kelamin perempuan atau laki-laki tanpa ketahapan selanjutnya yaitu pencarian nilai MFCC. Sehingga dalam penelitian ini dengan data latih dan data uji yang digunakan lebih baik menggunakan rentang frekuensi *overlapping* pada rentang 180 – 210 Hz untuk mengurangi kemungkinan kesalahan identifikasi yang terjadi pada sistem.

Pada penelitian selanjutnya, diharapkan penelitian ini bisa dikembangkan lagi dengan menggunakan data yang lebih besar, kemudian juga melakukan *tuning* parameter salah satu nya dengan melakukan analisa pengaruh jenis suara manusia (sopran, mezzo sopran, alto, tenor, barito, dan bass), dan juga melakukan analisa untuk meningkatkan akurasi pada perhitungan Mahalanobis Distance yang dapat mempengaruhi hasil klasifikasi yang lebih optimal sehingga mengurangi hasil kesalahan identifikasi.

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan dalam mengidentifikasi jenis kelamin secara *real-time* dengan perangkat Raspberry Pi, dan telah melalui pengujian dan analisa yang lebih mendalam. Maka dapat disimpulkan bahwa identifikasi jenis kelamin berdasarkan suara bisa dilakukan pada perangkat Raspberry Pi secara *real-time* dengan menggunakan parameter pitch dengan algoritma YIN dan parameter ekstraksi ciri MFCC. Pada penelitian ini, parameter range frekuensi *overlapping* berpengaruh pada hasil klasifikasi yang dihasilkan. Pada skema pertama pengujian dengan rentang pitch *overlapping* 150 – 210 Hz dan pada skema kedua menggunakan range frekuensi *overlapping* 180 – 210 Hz. Hasil pada skema kedua lebih baik dari pada skema pertama, pada data test perempuan di *closed test* yang semula 98% naik menjadi 100 %, kemudian pada *open test* yang semula 92% naik menjadi 96%. Sedangkan pada pengujian data test laki-laki di *closed test* yang semula 92% naik menjadi 98%, kemudian pada pengujian *open test* yang semula 90% naik menjadi 94%. Pada penelitian ini dengan data latih dan data uji yang telah digunakan lebih baik menggunakan rentang frekuensi *overlapping* pada rentang 180 – 210 Hz. Hal ini bertujuan untuk mengurangi kemungkinan kesalahan identifikasi dengan sistem yang diajukan. Dimana sistem tahapan pertama menggunakan parameter pitch yang apabila data suara termasuk dalam kategori *overlapping* akan mengalami tahapan kedua dengan menggunakan parameter ekstraksi ciri MFCC.

Daftar Pustaka

- [1] S.Chaudhary and D.K. Sharma, “Gender Identification based on Voice Signal Characteristics”, in International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2018.
- [2] Krishna D N, et all, “Language Independent Gender Identification From Raw Waveform Using Multi-Scale Convolutional Neural Networks”, in 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), ICASSP 2020.
- [3] B. Jolad and R. Khanai, “An Art of Speech Recognition: A Review” in 2nd International Conference on Signal Processing and Communication, ICSPC 2019 - Proceedings, 2019, pp. 31-35.
- [4] Furui, S., “Digital Speech Processing, Synthesis and Recognition”, Marcel Dekker Inc., New York. 2001.
- [5] A. P. Simpson, “Phonetic differences between male and female speech”, *Language and Linguistics Compass*, vol. 3, no. 2, pp. 621–640, 2009.
- [6] M. Alhussein, Z. Ali, M. Imran, and W. Abdul, “Automatic gender detection based on characteristics of vocal folds for mobile healthcare system” *Mobile Information Systems*, vol. 2016, 2016.
- [7] Y. Hu, D. Wu, and A. Nucci, “Pitch-based gender identification with two-stage classification,” *Security and Communication Networks*, vol. 5, no. 2, pp. 211–225, 2012.
- [8] K. Rakesh, S. Dutta, and K. Shama, “Gender recognition using speech processing techniques in labview”, *International Journal of Advances in Engineering & Technology*, vol. 1, no. 2, pp. 51–63, 2011.
- [9] M. Abouelenien, V. Perez-Rosas, R. Mihalcea, and M. Burzo, “Multi-modal gender detection” in *Proceedings of the 19th ACM International Conference on Multimodal Interaction*. ACM, 2017, pp. 302–311.
- [10] M. Kumari and I. Ali, “An efficient algorithm for gender detection using voice samples”, in *Communication, Control and Intelligent Systems (CCIS)*, 2015. IEEE, 2015, pp. 221–226.
- [11] A.de Cheveigne and H. Kawahara, “YIN, a fundamental frequency estimator for speech and music”, *J. Acoust. Soc. Am.*, Vol. 111, No. 4, April 2002.
- [12] F.Aldhi and M.Huda, “Sistem Klasifikasi Gender Berdasarkan Suara Pada Raspberry Pi,” *Jurnal Elektro PENS*, Vol.2, No.2, 2018.
- [13] Y. Al-Irhayim and A.Abdulkafor, “Speaker Gender Recognition Using Hidden Markov Model” *ISSN 2222-1719*, Vol.7, No.4, 2016.
- [14] Yakun Hu, Dapeng Wu and Antonio Nucci, “Pitch-based gender identification with two-stage classification”, in *Wiley Online Library SECURITY AND COMMUNICATION NETWORKS*, 2011.

- [15] Zizhu Fan et al, "Principal Component Analysis Integrating Mahalanobis Distance for Face Recognition", in 2013 Second International Conference on Robot, Vision and Signal Processing, 2013.
- [16] Chang-Young Lee, "A Study on the Optimal Mahalanobis Distance for Speech Recognition", Speech Science, Vol. 13 No. 4, 2006.