

# Klasifikasi Jenis Kelamin Berdasarkan *Pitch* Suara Menggunakan *Metode Pitch Detection Algorithm*

<sup>1</sup>Irham Surya Pratama, <sup>2</sup>Felix Indra Kurniadi

<sup>1,2</sup>Tanri Abeng University, School of Engineering and Technology, Informatics Engineering  
<sup>1</sup>irham.surya@student.tau.ac.id  
<sup>2</sup>felixindra@tau.ac.id

Diterima: 31 Agustus 2018  
Disetujui: 26 September 2018

**Abstract**—Gender classification is one of the major problems in speech analysis. Human voices exhibit very interesting aspects i.e. pitch, median, frequency etc. In this research we proposed an approach using pitch to classify each gender. Pitch Detection Algorithm were used as main method. This method gives 61% using Pitch Detection Algorithm

**Index Terms**—Gender, Pitch Detection Algorithm, Pitch

## I. PENDAHULUAN

Manusia dapat dikelompokkan berdasarkan jenis kelaminnya, yaitu pria dan wanita. Setiap manusia dengan manusia yang lain memiliki perbedaan karakteristik, salah satunya dapat dilihat dari suaranya. Tipe suara pada pria dibagi menjadi tenor, bariton, dan bass, sedangkan pada wanita dibagi menjadi sopran, mezzo sopran, dan alto, namun sebagian besar orang terutama orang awam mengalami kesulitan dalam menentukan tipe suara mereka [1]. Bila didengarkan dengan seksama, suara wanita cenderung lebih tinggi daripada suara pria [2].

Di antara perbedaan-perbedaan masih terdapat kemiripan antara suara pria satu dengan pria lain dan antara suara wanita satu dengan wanita lain. Dalam sekian banyak perbedaan yang ada pada masing-masing individu, ada beberapa kemiripan dari tiap-tiap kaum pria dan kaum wanita yang bisa diamati secara langsung. Salah satunya adalah suara mereka, walaupun suara masing-masing individu manusia berbeda-beda, namun ada kemiripan antara suara wanita yang satu dengan wanita yang lainnya, juga suara pria yang satu dan suara pria yang lainnya. Bila didengarkan dengan seksama, suara wanita cenderung lebih tinggi bila dibandingkan dengan suara pria. Melalui perbedaan tersebut, bias dilakukan identifikasi terhadap suara manusia melalui media komputer dengan memperhatikan perbandingan frekuensi suaranya, agar dapat digolongkan ke dalam suara pria atau wanita[3].

Penelitian ini dibuat dengan tujuan:

- ✓ Mengetahui tingkatan *Pitch* pada sinyal suara
- ✓ Melakukan klasifikasi suara berdasarkan *gender* (jenis kelamin)

Menghitung akurasi dari hasil pengklasifikasian suara. Manfaat yang dapat diambil dari penelitian ini adalah mendeteksi *pitch* suara dari setiap sampel dan mengklasifikasikan suara berdasarkan *gender* (jenis kelamin), sehingga dapat membuat komputer mampu mengenali suara manusia berdasarkan jenis kelamin.

Beberapa penelitian sebelumnya untuk melakukan pengklasifikasian jenis kelamin berdasarkan suara pernah dilakukan oleh Jain, penelitian ini menggunakan metode *Pitch Detection Method*, dan *average magnitude difference function* (AMDF) hasil yang didapatkan mencapai akurasi 99% dengan mengklasifikasi suara artis Internasional dan artis India dengan jumlah masing-masing 100 sampel di setiap jenis kelamin [4]. Harb menggunakan metode *Piecewise Gaussian Modeling* dan *Neural Network*, menyatakan bahwa *Pitch* merupakan bawaan setiap individu yang dapat menjadi pembeda setiap suara pria dan wanita. Estimasi suara dengan menggunakan *Pitch* dapat dioptimalkan dengan porsi suara bersih tanpa *noise*[5]. Rahuul mencoba mengatasi permasalahan klasifikasi gender ini dengan menggunakan beberapa metode machine learning, dimana dia hanya mengambil fitur dengan *mean*, *standard deviation* dan *median* yang kemudian diklasifikasikan menggunakan *Linear Discriminant Analysis* (LDA), K-Nearest

Neighbor (KNN), CART, *Random Forest* dan *Support Vector Machine* [6]. Berdasarkan penelitian-penelitian sebelumnya maka pada penelitian ini akan menggunakan metode *Pitch Detection Algorithm* untuk mengklasifikasi jenis kelamin berdasarkan *pitch* suara hal ini dikarenakan berdasarkan hasil yang dilakukan oleh [4], penggunaan metode PDA memberikan hasil yang mendekati sempurna .

Artikel Ilmiah ini dibagi menjadi lima bagian: **Bab I** yang menjelaskan tentang Pendahuluan, **Bab II** yang menjelaskan tentang Landasan Teori, **Bab III** menjelaskan tentang Metodologi, **Bab IV** menjelaskan tentang Hasil klasifikasi, dan **Bab V** menjelaskan tentang kesimpulan dari artikel ini.

## II. LANDASAN TEORI

### A. *Pitch (Nada)*

*Pitch* adalah tinggi rendah nada dalam suatu bunyian. *Pitch* sendiri berkaitan dengan getaran yang dihasilkan oleh instrument maupun suara manusia. Jika getarannya semakin banyak maka nada yang dihasilkannya pun semakin tinggi [7]. Suara yang tinggi identik dengan suara wanita, sedangkan suara rendah lebih identik kepada suara pria, oleh karena itu *pitch* dapat juga digunakan sebagai salah satu metode dalam mengklasifikasi.

### B. *Pitch Detection Algorithm*

*Pitch Detection Algorithm* adalah algoritma yang di desain untuk mengestimasi nada dasar dari sebuah Quasi Periodik atau Signal Virtual Periodik, yang biasanya merupakan rekaman digital dari suara maupun tangga nada musik [4]. Pada dasarnya *Pitch Detection Algorithm* menggunakan teknik short-term analisis [8]. Ekstraksi *Pitch* yang akurat sudah pernah di demonstrasikan dan dijadikan sebuah syarat yang penting dalam Speech Coding, Speech Compression, Speech Synthesis, Speech Recognition dan Speaker Identification [9].

### C. *Precision, Recall dan Specificity*

*Precision* berguna untuk melihat persentase prediksi data sebagai positif yang benar. *Recall* berguna untuk mengukur tingkat keberhasilan metode yang digunakan dalam menemukan sebuah informasi. Sedangkan *Specificity* berguna untuk mengetahui seberapa baik sebuah tes yang telah dilakukan dalam mencegah kesalahan yang ada. *Accuracy* adalah

ketepatan hasil klasifikasi dan fakta yang ada [10].

## III. METODOLOGI

Metodologi penelitian yang diterapkan pada penelitian ini dapat dilihat pada gambar 1 berikut:



Gambar 1. Metodologi Penelitian

### A. Pengumpulan Sampel Suara

Data sampel sejumlah 92 file suara berformat mp3 dan wav. Sampel suara terdiri dari 44 suara pria dan 48 suara wanita. Data diperoleh dari rekaman suara pada *Handphone* dengan durasi berkisar antara 3 – 5 detik dan dikirim melalui jejaring sosial. Isi rekaman sama, yaitu dengan menyebutkan kalimat “Halo, nama saya (nama panggilan)” dengan volume suara sedang dan tinggi sehingga diperoleh dua buah sampel data dalam sekali pengambilan sampel. Perekaman dilakukan dengan bantuan aplikasi *Audacity* dengan format PCM 8000Hz, 16 bit, 1 channel (mono) [11].

### B. *Konversi Suara*

Data yang telah direkam dikirim menggunakan jejaring sosial dan mengalami proses kompresi otomatis sehingga ukuran file suara lebih ringan daripada file asli. Data yang memiliki ekstensi file selain *.wav* mengalami konversi agar dapat diproses. Data sampel yang memiliki *noise* terlebih dahulu harus melalui tahap *noise reduction* agar hasil yang diperoleh lebih akurat.

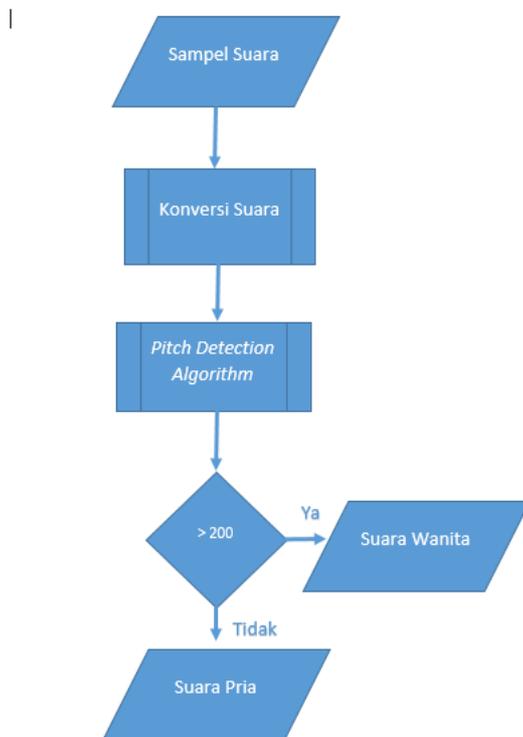
### C. Pencarian Pitch Suara

Dataset dikelompokkan menjadi dua, dan proses pencarian *Pitch* suara yang terlebih dahulu dilakukan adalah suara pria. Satu per satu dataset di inputkan kedalam perangkat lunak *Audacity* dan dengan *Tools* pendeteksi *Pitch* yang telah

tersedia dapat dilakukan pendeteksian *Pitch* dengan singkat.

#### D. Klasifikasi

Klasifikasi merupakan salah satu dari aplikasi proses *digital signal processing* (pengolahan sinyal digital) termasuk didalamnya *speech recognition*. Disini dianggap bahwa *pitch* yang melebihi dari 200 dianggap sebagai suara seorang wanita, hal ini sesuai dengan apa yang dilakukan oleh Jain Proses pengklasifikasian dilakukan satu per satu sesuai banyaknya sampel dan dilakukan berulang seperti proses berikut :



Gambar 2. Metodologi klasifikasi menggunakan Pitch Detection Algorithm

Angka yang diperoleh setelah proses klasifikasi dicatat dan dikalkulasi secara manual sehingga diperoleh hasil yang terstruktur.

#### E. Evaluasi Measurement

Proses evaluasi terhadap hasil yang diperoleh akan menggunakan akurasi, precision, recall dan specificity. Hal ini dilakukan untuk memberikan gambaran terhadap hasil dari nilai yang didapatkan dari metode PDA. *Precision* digunakan untuk menguji nilai peluang terhadap *classifier exactness*, sedangkan recall digunakan untuk menguji nilai peluang terhadap *classifier completeness* [12][10]. Sedangkan specificity untuk mengukur nilai kesalahan pada sebuah *classifier* [13].

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} * 100\% \quad (1)$$

$$precision = \frac{tp}{tp + fp} * 100\% \quad (2)$$

$$recall = \frac{tp}{tp + fn} * 100\% \quad (3)$$

$$specificity = \frac{tn}{(fp + tn)} \quad (4)$$

Dimana *tp* dan *tn* adalah prediksi yang sesuai dengan aslinya dan *fn* dan *fp* adalah prediksi yang tidak sesuai dengan aslinya.

## IV. HASIL

Berdasarkan proses yang telah dilakukan, hasil yang diperoleh adalah sebagai berikut :

Tabel 1. *Confussion Matrix* Hasil Klasifikasi Gender berdasarkan *Pitch* Suara

Hasil	Laki – laki	Perempuan
Laki – laki	33	11
Perempuan	25	23

Tabel 2. Evaluasi Hasil Klasifikasi Gender berdasarkan *Pitch* Suara

Evaluasi	Hasil
<i>Precision</i>	75%
<i>Recall</i>	56%
<i>Specificity</i>	67%
<i>Accuracy</i>	61%

Pada 44 sampel suara laki-laki, di dapat akurasi yang lebih tinggi dikarenakan banyak sampel suara yang memiliki volume suara yang pelan sehingga *Pitch* suara yang terdeteksi juga rendah, dan terklasifikasi sebagai suara laki-laki dikarenakan *Pitch* tidak melebihi angka 200Hz. Pada 48 sampel suara perempuan, di dapat akurasi yang rendah dikarenakan volume suara yang sangat kecil sehingga *Pitch* suara yang dihasilkan juga sangat rendah, hingga *Pitch* yang terdeteksi tidak mencapai angka 200Hz dan terklasifikasi sebagai suara laki-laki.

## V. KESIMPULAN

Dari seluruh metodologi yang telah dilakukan didapat akurasi yang belum cukup tinggi

dikarenakan pitch suara sangat dipengaruhi oleh volume suara, kualitas mikrofon, durasi rekaman, kompresi file yang disebabkan oleh karena proses perekaman yang dilakukan menggunakan *Tools* dari aplikasi jejaring sosial *Line*, dan noise yang masih terdapat pada beberapa rekaman. Volume yang sangat pelan akan mengeluarkan pitch minimum yaitu : 33.311 Hz meskipun suara yang di deteksi adalah suara wanita, namun hal ini menimbulkan akurasi yang tinggi pada saat proses klasifikasi pada suara laki-laki. Durasi rekaman berkisar antara 3 – 5 detik (tidak sama) menyebabkan akurasi berkurang. Kompresi file rekaman menyebabkan ukuran file menjadi lebih ringan dan hal tersebut merubah kualitas asli rekaman.

#### DAFTAR PUSTAKA

- [1] R. Dwifabrianti, B. Hidayat, and I. Wijayanto, "Analisis Deteksi Tipe Suara Pada Pria Dan Wanita Menggunakan Metode Mel-Frequency Cepstral Coefficient Dan Klasifikasi Support Vector Machine Multi-Kelas One-Against-All," 2013.
- [2] Mursyidah, Jamilah, and Zayya, "Pengenalan Karakter Suara Laki-Laki Aceh Menggunakan Metode FFT ( Fast Fourier Transform )," *J. Infomedia*, vol. 2, no. 1, pp. 20–24, 2017.
- [3] B. widi prasetya, B. Santoso, and J. Purwadi, "Identifikasi Suara Pria Dan Wanita Berdasarkan Frekuensi Suara," vol. 4, no. 2, pp. 2–9, 2008.
- [4] N. Jain and D. Kaushik, "Gender Voice Recognition Through Speech," India, 2014.
- [5] H. Harb and L. Chen, "Voice-based gender identification in multimedia applications," *J. Intell. Inf. Syst.*, vol. 24, no. 2–3, pp. 179–198, 2005.
- [6] A. Raahul, R. Sapthagiri, K. Pankaj, and V. Vijayarajan, "Voice based gender classification using machine learning," in *IOP Conference Series: Materials Science and Engineering* 263, 2017.
- [7] S. Endah and D. Mutiara KN, "KLASIFIKASI UCAPAN KATA DENGAN SUPPORT VECTOR MACHINE," *J. Masy. Inform.*, vol. 3, 2012.
- [8] L. Tan and M. Karnjanadecha, "Pitch detection algorithm: autocorrelation method and AMDF," in *3rd International Symposium on Communications and Information Technology*, 2003, pp. 551–556.
- [9] K. Abdullah-Al-Mamun, F. Sarker, and G. Muhammad, "A High Resolution Pitch Detection Algorithm Based on AMDF and ACF," *J. Sci. Res.*, vol. 1, no. 3, pp. 508–515, 2009.
- [10] J. Brownlee, "Classification Accuracy is not Enough Measures You Can See," 2014. [Online]. Available: <https://machinelearningmastery.com/classification-accuracy-is-not-enough-more-performance-measures-you-can-use/>. [Accessed: 06-Apr-2018].
- [11] R. Yessivirna, Marji, and D. E. Ratnawati, "Klasifikasi Suara Berdasarkan Gender (Jenis Kelamin) Dengan Metode K-Nearest Neighbor (Knn)," *J. Ilmu Komput.*, vol. 1, pp. 1–9, 2011.
- [12] C. Goutte and E. Gaussier, "A Probabilistic Interpretation of Precision , Recall and F -Score , with Implication for Evaluation," in *European Conference on Information Retrieval*, 2005, pp. 345–359.
- [13] R. Parikh, A. Mathai, S. Parikh, G. Chandra Sekhar, and R. Thomas, "Understanding and using sensitivity, specificity and predictive values," *Indian J. Ophthalmol.*, vol. 56, no. 1, pp. 45–50, Feb. 2008.