

Analisa Sinyal Suara untuk Deteksi Gangguan Pita Suara dengan Menggunakan Metode *Backpropagation Neural Network*

1st Shaina Diva
Prodi S1 Teknik Fisika
Fakultas Teknik Elektro
Bandung, Jawa Barat

shainadiva@student.telkomuniversity.a
c.id

2nd Amaliyah Rohsari Indah Utami
Prodi S1 Teknik Fisika
Fakultas Teknik Elektro
Bandung, Jawa Barat

amaliyahriu@telkomuniversity.ac.id

3rd Hertiana Bethaningtyas D.K
Prodi S1 Teknik Fisika
Fakultas Teknik Elektro
Bandung, Jawa Barat

hertiana@telkomuniversity.ac.id

Abstrak — Suara merupakan salah satu sarana komunikasi yang memegang peranan penting dalam penyampaian informasi, salah satunya suara yang dihasilkan oleh organ tubuh manusia yang berasal dari getaran pita suara. Sama seperti jaringan tubuh lain pita suara juga dapat mengalami kerusakan yang disebut penyakit kelainan pita suara. Penelitian ini ditujukan untuk mengembakan ragam metode non-invasif pada diagnosa kelainan pita suara, penelitian ini dilakukan dengan tujuan mengetahui arsitektur model yang optimal dan tingkat akurasi dari analisis untuk deteksi kelainan pita suara. Pre-processing dilakukan dengan normalisasi, *framing*, dan *windowing*, lalu dilakukan ekstraksi ciri dengan *Linear Predictive Coding* (LPC) dan *Mel Frequency Cepstral Coefficients* dimana masing masing fitur sebanyak 12 dan 13 lalu digabung menjadi satu vektor. lalu dilakukan Analisis dengan metode Backpropagation neural Network dengan dua data dengan variabel yang berbeda yakni data dengan neuron bervariasi dan tidak bervariasi hingga menghasilkan arsitektur optimal pada data dengan neuron tidak bervariasi dengan hidden layer 4 dan Epoch 150 menghasilkan akurasi 89,65 %, presisi dan recall sebesar 86,6% dan 92,8 %

Kata kunci— pita suara, LPC, MFCC, BNN, akurasi, presisi, recall

I. PENDAHULUAN

Berbicara merupakan salah satu media komunikasi verbal yang sangat efektif dalam kehidupan sehari-hari. Ketika berbicara suara tersebut berasal dari pita suara yang terletak pada tenggorokkan manusia, sama halnya seperti organ tubuh manusia lainnya pita suara juga dapat mengalami gangguan ataupun kerusakan. Untuk mendeteksi adanya kelainan pada pita suara diperlukan tenaga medis dan peralatan yang mampu menunjang diagnosa salah satunya yakni tindakan laringoskop atau stoboskopi yang bersifat *invasive* ke tenggorokan. Salah satu kelemahan dari tindakan tersebut yakni memerlukan biaya yang cukup mahal untuk mendapatkan hasil dignosanya. Sehingga diharapkan pada penelitian ini dapat membantu dalam perkembangan metode alternatif hingga mampu membantu tenaga medis dan tidak mengganggu kenyamanan pasien, salah satunya dengan metode *non-invasive*. [3]

Pada penelitian yang dilakukan oleh Ardi Wicaksono dkk yang berjudul "*Combining Itakura Saito Distance and Backpropagation Neural Network to Improve Sound Accuracy in Audio Forensic*" menggabungkan metode *speech processing* yaitu *itakura saito distance* dengan BNN yang digunakan untuk memperkuat tingkat akurasi dalam

mengidentifikasi suatu rekaman suara, akurasinya diukur dari nilai kedekatan frekuensi dengan *spectrum* antara Suara Asli dengan pembanding menghasilkan akurasi yang tertinggi yang identik lebih dari 95% [7].

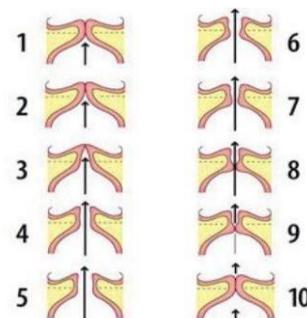
Berdasarkan latar belakang tersebut, penelitian yang direncanakan oleh peneliti menggunakan metode *backpropagation neural network* sebagai bentuk pengembangan metode *non-invasive* yang sudah ada dari segi akurasi model yang dibangun. Dengan adanya penelitian ini, peneliti berharap metode yang digunakan pada penelitian ini mampu membantu para ahli dalam melakukan diagnosa terhadap pasien yang menderita kelainan pita suara.

II. KAJIAN TEORI

A. Proses Pembentukan Suara

Proses terjadinya getaran pada pita suara, yang tentunya akan menghasilkan suara, awal mula Tekanan udara bergerak ke atas pita suara yang berada dalam posisi tertentu, sehingga tekanan udara mampu membuka lapisan getar bagian bawah dari pita suara, posisi pita suara berada dalam keadaan posisi tetap, kemudian tekanan akan terus bergerak ke atas sehingga bagian atas dari pita suara akan terbuka, sehingga tekanan yang lemah tercipta di balik kolom udara yang beregrak dang menghasilkan efek Bernoulli yang akan menyebabkan bagian bawah menutup diikuti oleh bagian atas, penutupan pita suara. [8] berikut merupakan gambar proses terjadinya getaran pada pita suara.

Gambar 2.1 Proses terjadinya getaran pada pita suara [8]



B. Gangguan Pita Suara

Perubahan suara biasanya sangat berkaitan dengan gangguan pita suara, gejala perubahan kualitas suara dapat

dirasakan diantaranya yaitu suara rendah dan Serak. Orang yang mengalami masalah suara biasanya mengeluh kehilangan suara, kehilangan daya tahan tubuh, dan terkadang terasa nyeri di tenggorokan ketika sedang berbicara. Gangguan pita suara pada umumnya dapat meliputi, seperti: Nodul, polip, kista, laringitis, tumor, paresis, bahkan hingga kanker tenggorokkan.

Gambar 2.2 Polip pada pita suara dan kanker Tenggorokan [2]



III. METODE

A. Linear Predictive Coding (LPC) dan Mel-Frequency Cepstral Coefficients (MFCC)

Linear prediction coding adalah salah satu metode analisis sinyal yang secara matematis mudah diterapkan untuk ekstraksi fitur dalam pengenalan kata. LPC juga dapat digunakan sebagai perkiraan kemungkinan, *Linear prediction* sebagaimana mengoptimalkan kecepatan bit yang rendah dengan kualitas suara yang memadai. LPC memiliki beberapa kelebihan salah satunya yakni komputasi yang dibutuhkan lebih sedikit dibandingkan analisis lainnya yang umumnya digunakan dalam *speech recognition*.

MFCC (*Mel-Frequency Cepstral Coefficients*) merupakan salah satu metode yang biasa digunakan untuk pemrosesan suara, terutama untuk pengenalan ucapan dan analisis audio. MFCC bekerja dengan mengekstraksi fitur penting dari sinyal suara, mirip dengan bagaimana manusia mendengar dan membedakan suara. Sederhananya, MFCC membantu komputer memahami pola suara dengan mengonversi gelombang suara ke bentuk yang mudah dianalisis. Dalam prosesnya MFCC melibatkan Pemecahan suara menjadi potongan kecil yang bisa dilakukan dengan *Framing* dan *Windowing Audio*, lalu mentransformasi ke Frekuensi menggunakan FFT, Menerapkan skala Mel yang mencerminkan cara manusia merespon frekuensi, lalu Transformasi Cepstral yang membantu mengekstraksi fitur utama dari suara

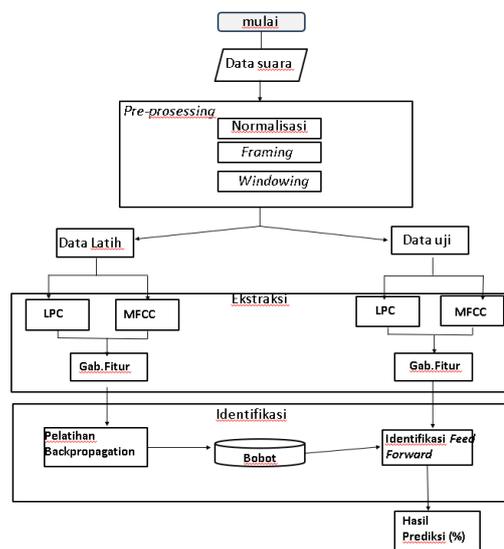
B. Backpropagation Neural Network

Jaringan saraf tiruan (JST) merupakan bentuk dari representasi yang dibuat oleh manusia untuk mensimulasikan proses pembelajaran di otak manusia. Pendekatan dengan jaringan saraf tiruan bisa meniru setiap hubungan yang kompleks dan *non-Linear (Neuron)* dan banyak digunakan pada daerah *forecasting*. *Forecasting* dasarnya merupakan suatu anggapan tentang sesuatu yang terjadi pada suatu peristiwa atau kejadian yang akan datang. Jaringan saraf tiruan memiliki beberapa struktur jaringan yaitu *single layer*

network, multilayer network, model JST dua dengan *feedback*, dan Model JST lapisan kompetitif.[7]

Salah satu sub bidang kecerdasan buatan yang dapat diandalkan dalam memprediksi ialah *Backpropagation Neural Network (BNN)*. BNN merupakan algoritma yang efektif dalam memberikan solusi penurunan gradien yang dapat meminimalkan kuadrat *input* yang *error*. Metode BNN sendiri memiliki tiga tahapan utama, Adapun tiga tahapan tersebut tahapan perambatan maju, tahapan perambatan balik, dan tahapan perubahan bobot dan bias. Algoritma pada BNN sering dikatakan sebagai algoritma *multilayer* yang memiliki parameter diantaranya *input layer, hidden layer, dan output layer*. [5]

C. ALUR PENGOLAHAN DATA



Gambar 2.3 Diagram Pengolahan Data

IV. HASIL DAN PEMBAHASAN

Pada penelitian ini digunakan sebanyak 145 data rekaman suara dengan bentuk *.wav*, data suara dibagi menjadi dua, yakni kelompok suara individu sehat sebanyak 79 dan kelompok suara pasien sakit sebanyak 68. Pada penelitian ini pada tahapan *pre-processing* dilakukan normalisasi, *framing*, dan *windowing*. Hingga ekstraksi ciri LPC dan MFCC yang dilakukan secara bertahap dengan 12 fitur dan 13 fitur hingga digabung menjadi satu vektor berukuran 25 yang menjadi *input* pada model BNN. Serta pada pemodelan BNN dibagi menjadi dua berdasarkan neuron yang digunakan dengan *hidden layer* yang sama pada setiap percobaan. Berikut merupakan tabel dari data dengan neuron variatif dan data dengan neuron tidak variatif.

Tabel 4.1. Data dengan Neuron Variatif

Jumlah Hidden Layer	Struktur Neuron	Nilai Epoch	Loss Latih	Loss Validasi	Akurasi (%)
2	16	50	0,6465	0,5868	69

	32	100	0,4954	0,5225	76
		150	0,4723	0,4768	76
3	16	50	0,5999	0,6022	66
	32	100	0,5524	0,618	72
	64		0,4995	0,5114	83
4	16	50	0,5782	0,5857	76
	32	100	0,5091	0,5464	72
	64		150	0,4671	0,4518
	128				
5	16	50	0,605	0,6222	72
	32	100	0,5464	0,5745	72
	64		150	0,454	0,5335
	128				
	256				
6	16	50	0,5605	0,5903	72
	32	100	0,522	0,5811	72
	64				
	128	150	0,4951	0,5284	72
	256				
512					

	64	100	0,3974	0,4885	76
		150	0,3788	0,4636	79
3	32	50	0,5565	0,5429	69
	64	100	0,4563	0,5351	79
	64		150	0,3788	0,4793
4	32	50	0,5961	0,6085	79
	64	100	0,4616	0,558	72
	64		150	0,3724	0,55405
	32				
5	32	50	0,6281	0,6132	72
	64	100	0,4959	0,5259	79
	64		150	0,3581	0,507
	32				
	32				
6	32	50	0,6355	0,6311	72
	64	100	0,5455	0,5748	86
	64		150	0,4147	0,5068
	32				
	32				

Pada data variatif menggunakan data sinyal suara pada penelitian ini dapat disimpulkan model semakin baik dalam belajar dan mengidentifikasi dengan baik pada saat jumlah hidden layer 4 dengan nilai epoch 150 dimana akurasi validasi sebesar 86% sehingga menunjukkan semakin bertambah hidden layer kapasitas representasional model bertambah. Namun terlalu banyak jumlah hidden layer dan nilai neuron yang tinggi juga tidak berarti dapat meningkatkan kinerja model karena dapat meningkatkan resiko *overfitting*, waktu komputasi yang lebih lama, dan masalah lainnya seperti *vanishing/exploding gradients* selama pelatihan.

Tabel 3.2. Data dengan Neuron tidak variatif

Jumlah Hidden Layer	Struktur Neuron	Nilai Epoch	Loss Latih	Loss Validasi	Akurasi (%)
2	32	50	0,4706	0,5275	72

Jadi performa paling optimal pada penelitian ini ialah pada *hidden layer* 4 dengan waktu pelatihan selama 150 *epoch* dengan jumlah neuron yang tidak bervariasi yakni 32 dan 64 dengan data suara yang digunakan pada penelitian ini sehingga dapat dikatakan dengan *hidden layer* dan jumlah neuron pada penelitian ini berada dalam kondisi *goodfitting* dan tidak mengalami *underfitting/overfitting*. kemungkinan untuk perubahan arsitektur yang paling optimal bisa berubah jika data bertambah atau berkurang dapat disesuaikan dengan data yang digunakan di masa depan

V. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan mengenai klasifikasi sinyal suara untuk deteksi gangguan pada pita suara menggunakan metode *Backpropagation Neural Network* (BPNN) dapat ditarik kesimpulan sebagai berikut:

1. Arsitektur Model yang paling optimal pada penelitian ini yakni: Melalui serangkaian eksperimen dengan arsitektur neuron variatif dan non-variatif, ditemukan bahwa kinerja model terbaik dicapai menggunakan arsitektur neuron tidak variatif dengan 4 *hidden layer* (menggunakan neuron 32 dan 64) dan pelatihan selama 150 *epoch*.

2. Model *Backpropagation Neural Network* dengan arsitektur optimal berhasil mencapai tingkat performa yang tinggi dalam mengklasifikasikan data uji, dengan rincian sebagai berikut:
- o Akurasi sebesar 89,55%, yang menunjukkan bahwa model mampu mengklasifikasikan hampir 90% data dengan benar.
 - o Presisi sebesar 87,87%, menunjukkan tingginya tingkat kebenaran dari prediksi yang dibuat oleh model.
 - o Recall sebesar 90,65%, menunjukkan kemampuan model yang sangat baik dalam mengidentifikasi seluruh kasus positif yang sebenarnya dari total data.

REFERENSI

- [1] Tim Redaksi KBBI. 2020. *"Kamus Besar Bahasa Indonesia"*. Badan Pengembangan dan Pembinaan Bahasa. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia, Jakarta.
- [2] Natthu Dhoke, Trupti. 2016. *Review of vocal cord analysis using speech for various voice disorders*. International Journal of Advanced Research in Computer and Communication Engineering. Vol 5. India.
- [3] Nurmalida, Firda dkk. 2012. *Analisa Deteksi Kelainan Pita Suara dengan Menggunakan Transformasi Wavelet Biorthogonal*. Universitas Telkom. Bandung.
- [4] M. Nasser, Ibrahim. 2019. *Lung Cancer Detection Using Artificial Neural Network*. Al-Azhar University. Gaza, Palestine.
- [5] Dilip Roy Chowdhury, dkk. 2011. *An Artificial Neural Network Model for Neonatal Disease Diagnosis*. International Journal of Artificial Intelligence and Expert Systems (IJAE), Volume (2): Issue (3).
- [6] Al-Milli, Nabeel. 2013. *Backpropagation Neural Network for Prediction of Heart Disease*. Zarga University College. Al-Balqa.
- [7] Nur. Nurul Izzah Luthfiah. 2021. *Analisis Sinyal Suara untuk Deteksi Penyakit Kelainan Suara dengan Convolutional Neural Network (CNN)*. Fakultas Teknik Elektro, Universitas Telkom.
- [8] Wicaksono, Ardy. 2020. *Penggabungan Metode Itakura Saito Distance dengan Backpropagation Neural Network untuk Peningkatan Akurasi Suara Pada audio Forensik*. Universitas Islam Indonesia. Yogyakarta
- [9] Syamsudin dkk, 2018. *Anatomi Suara, Anatomi Suara Kajian Fisika Medik*. Cetakan Pertama. Airlangga University. Surabaya.
- [10] Farida, Fitri. 2017. *Optimasi Lowpass Filter Mikrostrip Frekuensi 10,6 GHz dengan Metode Step-Impedansi*. Universitas Maritim Raja Ali Haji. Tanjung Pinang.
- [11] Jaehoon, Lee dkk. 2018. *Deep Neural Networks As Gaussian Processes*. Published as a conference paper at ICLR.
- [12] Azmi, Yamina dkk. 2018. *Identifikasi Emosi Dari Sinyal Suara Secara Real Time Menggunakan Linear Predictive Coding dan Backpropagation*. Universitas Jenderal Achmad Yani, Cimahi.
- [13] Rochman, Eka Mala Sari dkk. 2024. *Classification of hypertension disease using Artificial Neural Network (ANN) backpropagation method, case study in mitigating health risk: UPT Modopuro Mojokerto Health Center*. INTI International University. Negeri Sembilan, Malaysia.