Evaluasi Kinerja Plugin AI Clarity Vx dalam Mereduksi Microphone Bleeding pada Multitrack Vocal Live Sound Berdasarkan SNR

Alberthino Ramadhan Oktaviano[1*], Nurchim[2], Pramono[3]

[1]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa
[1*]220103172@mhs.udb.ac.id

[2]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa
[2]nurchim@udb.ac.id

[3]Teknik Informatika/Fakultas Ilmu Komputer
Universitas Duta Bangsa
[3]pramono@udb.ac.id

Abstrak— **Menjaga kejernihan sinyal vokal pada sistem suara pertunjukan langsung (*live sound*) sering kali terkendala oleh fenomena *microphone bleeding*, terutama akibat instrumen berdinamika keras. Keterbatasan perangkat konvensional dalam menyaring kebocoran tersebut mendorong pemanfaatan algoritma kecerdasan buatan melalui *plugin* Waves Clarity Vx. Penelitian ini bertujuan untuk mengevaluasi kinerja *plugin* tersebut dalam mereduksi *microphone bleeding* pada kanal vokal *multitrack* guna menentukan batas optimal terhadap intensitas pemrosesan. Dengan menggunakan pendekatan kuantitatif eksperimental melalui simulasi *Virtual Soundcheck*, pengujian dilakukan pada 15 sampel vokal dengan menerapkan variasi tingkat reduksi mulai dari 0%, 50%, 75%, hingga 100%. Tingkat kejernihan sinyal dievaluasi secara objektif berdasarkan pengukuran nilai amplitudo *Root Mean Square* (RMS) yang dikonversi menjadi *Signal-to-Noise Ratio* (SNR). Hasil analisis menunjukkan bahwa intensitas pemrosesan 50% adalah sebagai titik optimal karena mampu meredam kebisingan latar secara signifikan tanpa mengurangi kualitas sinyal vokal, sedangkan intensitas 100% terbukti stagnan dan justru berisiko mereduksi energi vokal utama. Dengan demikian, *plugin* AI Clarity Vx dinilai efisien untuk diimplementasikan secara langsung pada sistem *live sound* selama pengaturannya dikunci pada batas aman 50%.**

Abstract— **Maintaining the clarity of vocal signals in the *live sound* system is often hampered by the phenomenon of *microphone bleeding*, especially due to loud dynamic instruments. The limitations of conventional devices in filtering for leaks encourage the use of artificial intelligence algorithms through the Waves Clarity Vx *plugin*. Using an experimental quantitative approach through *Virtual Soundcheck* simulation, tests were performed on 15 vocal samples by applying variations in reduction rates ranging from 0%, 50%, 75%, to 100%. The level of signal clarity is objectively evaluated based on the measurement of the *Root Mean Square* (RMS) amplitude value converted to *the Signal-to-Noise Ratio* (SNR). The results of the analysis show that the processing intensity of 50% is the optimal point because it is able to significantly reduce background noise without reducing the quality of the vocal signal, while the intensity of 100% is proven to be stagnant and actually risks reducing the main vocal energy. As such, the Clarity Vx AI *plugin* is rated efficient to implement directly on the *live sound system* as long as the settings are locked to the 50% safe limit.**

Keywords— Artificial Intelligence, Clarity Vx, Live Sound, Microphone Bleeding, Signal-to-Noise Ratio.

I. PENDAHULUAN

Dalam dunia live sound menjaga kejernihan sinyal vokal merupakan hal yang sangat penting karena vokal menjadi fokus utama dari sebuah pertunjukan. Namun, kondisi di panggung yang dinamis sering kali memicu terjadinya microphone bleeding atau kebocoran suara [1]. Kebocoran ini terjadi ketika sinyal suara dari instrumen latar ikut tertangkap oleh mikrofon vokal [2]. Fenomena tersebut menyebabkan penurunan tingkat kejernihan vokal, terjadinya penumpukan frekuensi yang mengganggu, serta meningkatkan risiko umpan balik (*feedback*) saat terjadinya kenaikan volume kanal (*gain*).

Tantangan kebocoran suara ini menjadi lebih tinggi pada pertunjukan musik yang menggunakan instrumen jenis perkusi dan metalofon [3]. Instrumen tersebut memiliki karakteristik dinamika pukulan yang keras, sehingga gelombang suara rentan mendominasi area tangkap mikrofon vokal [4], [5]. Untuk mengurangi dampak kebocoran tersebut, perangkat pengolah suara konvensional seperti *noise gate* sering digunakan sebagai solusi standar. Namun, noise gate bekerja secara linear berdasarkan batas volume (*threshold*) [6]. Ketika vokalis mulai bernyanyi, gerbang pemroses akan terbuka sehingga kebocoran

suara instrumen yang keras tetap masuk ke saluran vokal tanpa adanya proses penyaringan.

Sebagai solusi dari keterbatasan metode konvensional tersebut, teknologi pemrosesan audio saat ini memanfaatkan algoritma kecerdasan buatan (AI) berbasis *Deep Neural Network* [7], [8], salah satunya adalah *plugin* Waves Clarity Vx. *Plugin* ini dirancang untuk mengenali karakteristik suara manusia dan memisahkan dari gangguan suara latar belakang. Penelitian-penelitian terkini membuktikan bahwa penggunaan AI pada plugin audio sangat efektif dalam mengotomatisasi pemrosesan multitrack [9], mampu mengisolasi vokal dari instrumen panggung dengan akurasi yang jauh melebihi metode pemotongan frekuensi tradisional [10], [11], serta dapat meredam derau (*noise suppression*) secara tangkas pada ekosistem audio waktu nyata [12]. Walau demikian, kinerja dari pemrosesan AI ini perlu dievaluasi secara objektif. Pengaturan parameter intensitas pemrosesan yang terlalu tinggi atau berlebih berisiko memotong elemen komponen frekuensi dari vokal utama itu sendiri. Oleh karena itu, pengujian yang terukur diperlukan melalui perhitungan nilai *Root Mean Square* (RMS) dan *Signal-to-Noise Ratio* (SNR) untuk mengetahui batasan efektivitas pemrosesan sinyal.

Penelitian ini dilakukan untuk mengevaluasi parameter intensitas pemrosesan *plugin* tersebut dalam mereduksi kebocoran sinyal pada saluran vokal. Pengujian dijalankan menggunakan metode simulasi (*Virtual Soundcheck*) dari data rekaman *multitrack* untuk mencari batas pengaturan parameter yang paling optimal. Melalui analisis tren data kuantitatif yang dihasilkan, hasil evaluasi dari pengujian simulasi ini nantinya ditujukan sebagai dasar acuan untuk mengetahui tingkat efisiensi *plugin* apabila diimplementasikan secara langsung pada sistem *live sound* di konser pangggung yang sebenarnya.

II. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif eksperimental yang dirancang secara sistematis melalui empat tahapan utama. Adapun tahapan alur penelitian dilaksanakan sebagai berikut:

Gambar 1. Tahap Penelitian

A. *Persiapan*

Tahapan persiapan menguraikan prosedur pengumpulan data primer serta penetapan spesifikasi perangkat pendukung yang digunakan sebagai media pengujian dalam penelitian.

1. Pengumpulan data

Sumber data yang digunakan dalam penelitian ini adalah data primer berupa rekaman *multitrack live sound* yang telah direkam dari beberapa event berbeda. Rekaman diperoleh menggunakan sistem *digital mixer* dan direkam dalam format *multitrack* untuk memungkinkan analisis terpisah pada kanal mikrofon vokal. Metode pengumpulan data dalam penelitian ini dilakukan melalui beberapa tahapan. Tahap pertama adalah mengumpulkan file *multitrack* hasil rekaman live sound yang tersedia. Selanjutnya, dilakukan proses seleksi data berdasarkan beberapa kriteria, yaitu kanal vokal yang mengandung *microphone bleeding* secara jelas, kondisi audio yang tidak mengalami *clipping* dan distorsi.

Jumlah sampel yang digunakan dalam penelitian ini adalah 15 sampel vokal dari 3 genre musik yaitu Campursari, Dangdut dan Karawitan. Setelah sampel yang sesuai diperoleh, tahap berikutnya adalah mengidentifikasi dua jenis segmen pada setiap sampel, yaitu segmen vokal dominan yang berperan sebagai *signal* utama dan segmen tanpa vokal aktif yang didominasi oleh kebocoran suara instrumen panggung (*microphone bleeding*) sebagai *noise* dengan durasi 60 detik. Data yang telah melalui proses seleksi dan identifikasi tersebut kemudian digunakan sebagai bahan eksperimen untuk menguji kinerja metode reduksi *microphone bleeding* yang diterapkan dalam penelitian ini.

2. Kebutuhan Perangkat

Penelitian ini didukung menggunakan tiga perangkat lunak utama untuk kebutuhan eksekusi dan analisis data yaitu:

a. PreSonus Studio One

Berfungsi sebagai *Digital Audio Workstation* (DAW) utama yang berperan sebagai media untuk simulasi *Virtual Soundcheck* dalam pemrosesan sinyal audio.

b. Waves Clarity Vx

Plugin audio berbasis *Artificial Intelligence* (AI) dengan arsitektur jaringan *Deep Neural Network* yang bertindak sebagai variabel pemroses inti untuk mengeksekusi reduksi kebocoran sinyal (*microphone bleeding*).

c. iZotope RX 11

Perangkat lunak analisis spektrum tingkat lanjut yang digunakan secara khusus pada fase pengukuran pasca-eksperimen guna mengekstrak data statistik gelombang (*wave statistics*) secara objektif.

B. Eksperimen

Tahap ini menjelaskan prosedur pengujian performa kinerja *plugin* Clarity Vx dalam mereduksi microphone bleeding pada 15 sampel vokal yang telah diseleksi.

1. Konfigurasi DAW Studio One

Eksperimen dijalankan menggunakan metode *Virtual Soundcheck* di dalam *Digital Audio Workstation* (DAW) Studio One. Berkas WAV vokal mentah dengan durasi 60 detik diimpor ke dalam audio *track* baru. Pada slot insert utama di setiap saluran, disisipkan *plugin* AI (*Artificial Intelligence*) Waves Clarity Vx sebagai pemroses tunggal untuk mereduksi pemisahan komponen suara.

Gambar 2. DAW Studio One

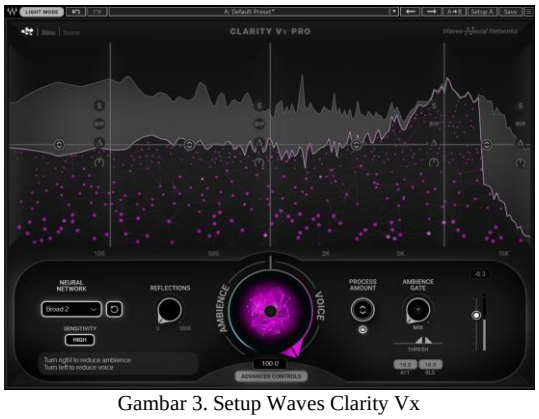
2. Penerapan Variasi Parameter Pemrosesan

a. Konfigurasi Plugin Clarity Vx

Parameter pertama yang disetting pada *plugin* ini yaitu Main Control (*Ambience vs Voice*) diset pada nilai maksimum (100%) dan dijaga konstan selama proses eksperimen. Hal ini bertujuan agar seluruh hasil pemrosesan merepresentasikan keluaran penuh dari *neural network*, sehingga perubahan kualitas sinyal hanya dipengaruhi oleh variasi intensitas parameter uji. Selain itu, konfigurasi berikutnya adalah:

Neural Network : *Broad* 2
Sensitivity : *High*

Pemilihan *Neural Network Broad* 2 didasarkan pada kemampuannya dalam memfokuskan pemisahan suara utama (vokal) dari suara latar dan sumber lain. Parameter Sensitivity diatur pada nilai High untuk meningkatkan kemampuan analisis sinyal pada kondisi dengan rasio *signal-to-noise* yang rendah, yang umumnya terjadi pada sistem live sound [13]. Contoh setup *plugin* Waves Clarity Vx dapat dilihat pada Gambar 3.



Gambar 3. Setup Waves Clarity Vx

Seluruh parameter tersebut dijaga konstan selama proses eksperimen untuk memastikan konsistensi hasil pengujian. Selama proses eksperimen, tidak digunakan efek tambahan seperti *equalizer* (EQ), *gate*, *compressor*, maupun efek lainnya. Hal ini bertujuan untuk menjaga variabel kontrol agar perubahan yang terjadi pada sinyal audio murni disebabkan oleh pemrosesan *plugin* Clarity Vx.

b. Proses Reduksi

Variasi tingkat reduksi ditentukan berdasarkan parameter *Process Amount*, yang berfungsi sebagai pengontrol utama intensitas

pemrosesan *neural network* dalam mereduksi *noise* atau *microphone bleeding*. Pada tahap ini, langkah pertama dilakukan dengan menerapkan nilai *Process Amount* 50%, 75% dan 100% pada data awal 0% (*baseline*) disetiap sampel track vokal yang kemudian di *ekspor* satu persatu untuk masing-masing *genre*. Setiap proses dilakukan pada segmen audio yang sama untuk memastikan konsistensi perbandingan antar variasi tingkat reduksi. Hasil dari setiap proses kemudian disimpan sebagai data output untuk tahap pengukuran dan analisis selanjutnya

3. Pengukuran

Metode pengukuran pada penelitian ini dilakukan dua tahap, yaitu nilai *Root Mean Square* (RMS) dan *Signal-to-Noise Ratio* (SNR) terhadap hasil dari proses reduksi yang dilakukan sebelumnya. Adapun langkah-langkah untuk pengukuran yang dilakukan adalah sebagai berikut:

a. Mengukur *Root Mean Square* (RMS)

Proses pengukuran nilai RMS dilakukan menggunakan perangkat lunak iZotope RX 11. Pada setiap sampel berdurasi 60 detik tersebut, dilakukan penentuan dua rentang waktu analitis yang presisi dengan variasi durasi sekitar 2 sampai 7 detik pada segmen *bleeding* dan vokal aktif. Segmen pertama adalah Segmen (*bleeding*), yaitu durasi spesifik saat vokalis sedang diam atau berhenti bernyanyi, di mana mikrofon vokal murni hanya menangkap suara kebocoran instrumen panggung. Segmen kedua adalah vokal aktif, yaitu durasi spesifik saat vokalis aktif bernyanyi secara dominan. Berikut penggunaan iZotope RX 11 untuk pengukuran RMS melalui fitur *Waveform Statistics* dapat dilihat pada Gambar 4.

Gambar 4. Penggunaan iZotope RX 11

Melalui fitur *Waveform Statistics*, nilai *Root Mean Square* (RMS) di ambil dari baris (Total RMS level) yang kemudian dicatat dalam satuan *desibel* (dB) dari kedua segmen tersebut. Nilai RMS yang diperoleh kemudian dikonversi secara matematis ke dalam nilai *Signal-to-Noise Ratio* (SNR) untuk menentukan tingkat kebersihan vokal.

b. Perhitungan *Signal-to-Noise Ratio* (SNR)

Nilai *Root Mean Square* (RMS) yang didapatkan dari tahap pengukuran sebelumnya dikonversi menjadi nilai *Signal-to-Noise* Ratio (SNR). Kalkulasi SNR didasarkan pada selisih logaritmik antara nilai RMS segmen vokal aktif (*signal*) dan segmen RMS *bleeding* (noise) [14]. Perhitungan dinyatakan menggunakan persamaan matematis sebagai berikut:

$$SNR_{(dB)} = RMS_{signal(dB)} - RMS_{noise(dB)}$$

Keterangan:

SNR : Rasio vokal terhadap kebocoran
RMS_*signal* : Nilai segmen vokal aktif
RMS_*noise* : Nilai segmen *noise/bleeding*

4. Analisis

Setelah nilai *Signal-to-Noise Ratio* (SNR) diperoleh dari konversi *Root Mean Square* (RMS), pengolahan data dilanjutkan menggunakan pendekatan statistik deskriptif komparatif. Langkah pertama adalah melakukan kalkulasi nilai rata-rata (*mean)* SNR dari kondisi awal (*baseline* 0%) ke setiap tingkat pemrosesan (50%, 75%, dan 100%). Pengelompokan nilai rata-rata ini dilakukan secara spesifik berdasarkan tiga *genre* musik (Campursari, Dangdut, dan Karawitan) guna memetakan tren pergerakan rasio kejernihan pada masing-masing *genre*.

Seluruh akumulasi data tersebut kemudian divisualisasikan ke dalam bentuk grafik garis. Prosedur analisis akhir difokuskan untuk mengevaluasi secara objektif tren pergerakan garis pada grafik tersebut guna menentukan batas optimal dari kinerja *plugin* Clarity Vx. Evaluasi berbasis tren grafik ini menjadi landasan utama untuk menarik kesimpulan mengenai batas maksimal intensitas pemrosesan yang dapat diterapkan tanpa mereduksi energi dari sinyal vokal utama.

III. HASIL DAN PEMBAHASAN

A. *Hasil Pengukuran Root Mean Square (RMS)*

Pengukuran dilakukan menggunakan perangkat lunak iZotope RX 11 untuk mengekstrak nilai RMS pada dua segmen analitis, yaitu segmen vokal aktif (*signal*) dan segmen *bleeding* (*noise*). Nilai RMS ini pada dasarnya mewakili tingkat kekerasan suara rata-rata pada rentang waktu tertentu. Sesuai dengan langkah kerja yang telah direncanakan, pengukuran dipisah menjadi dua kondisi, segmen vokal aktif (saat penyanyi sedang bernyanyi) dan segmen *bleeding* (saat penyanyi diam dan mikrofon hanya menangkap bocoran suara latar panggung). Hasil ekstraksi nilai RMS dari seluruh sampel pada ketiga *genre* dengan tingkat reduksi (0%, 50%, 75%, dan 100%) disajikan selengkapnya pada Tabel 1, Tabel 2, dan Tabel 3 berikut.

Tabel 1. RMS Vokal Campursari

Sampel	Segmen Durasi	RMS 0%	RMS 50%	RMS 75%	RMS 100%
Vokal 1	Bleeding 27,657-31,657	-38,07 dB	-48,68 dB	-48,54 dB	-48,46 dB
	Vokal aktif 23,649-27,649	-15,85 dB	-18,90 dB	-18,89 dB	-18,89 dB
Vokal 2	Bleeding 11,330-18,330	-32,16 dB	-56,31 dB	-56,32 dB	-56,32 dB
	Vokal aktif 04,423-11,423	-18,81 dB	-25,02 dB	-24,97 dB	-24,95 dB
Vokal 3	Bleeding 28,567-31,567	-30,49 dB	-63,37 dB	-64,22 dB	-64,75 dB
	Vokal aktif 31,536-34,536	-16,37 dB	-19,95 dB	-19,91 dB	-19,89 dB
Vokal 4	Bleeding 27,805-34,805	-31,01 dB	-48,67 dB	-48,57 dB	-48,52 dB
	Vokal aktif 20,805-27,805	-23,70 dB	-27,71 dB	-27,65 dB	-27,62 dB
Vokal 5	Bleeding 37,442-41,442	-31,53 dB	-42,48 dB	-42,35 dB	-42,27 dB
	Vokal aktif 34,155-38,155	-13,18 dB	-16,26 dB	-16,25 dB	-16,24 dB

Tabel 2. RMS Vokal Dangdut

Sampel	Segmen Durasi	RMS 0%	RMS 50%	RMS 75%	RMS 100%
Vokal 1	Bleeding 19,230-22,123	-32,16 dB	-44,95 dB	-44,85 dB	-44,81 dB
	Vokal aktif 22,153-25,046	-22,16 dB	-25,65 dB	-25,58 dB	-25,54 dB
Vokal 2	Bleeding 13,550-15,550	-37,07 dB	-51,69 dB	-51,75 dB	-51,77 dB
	Vokal aktif 15,615-17,615	-18,11 dB	-21,20 dB	-21,18 dB	-21,17 dB
Vokal 3	Bleeding 10.797-14,000	-26,82 dB	-64,87 dB	-65,55 dB	-65,85 dB
	Vokal aktif 14,000-17,203	-16,37 dB	-19,17 dB	-19,68 dB	-19,67 dB
Vokal 4	Bleeding 25,199-29,199	-32,74 dB	-44,54 dB	-44,39 dB	-44,30 dB
	Vokal aktif 30,051-34,584	-17,18 dB	-20,39 dB	-20,34 dB	-20,22 dB
Vokal 5	Bleeding 25,960-30,780	-24,65 dB	-40,50 dB	-40,38 dB	-40,30 dB
	Vokal aktif 37,273-42,092	-14,69 dB	-18,19 dB	-18,12 dB	-18,08 dB

Tabel 3. RMS Vokal Karawitan

Sampel	Segmen Durasi	RMS 0%	RMS 50%	RMS 75%	RMS 100%
Vokal 1	Bleeding 23,482-26,289	-36,23 dB	-63,56 dB	-63,58 dB	-63,56 dB
	Vokal aktif 20,693-23,500	-23,81 dB	-27,15 dB	-27,10 dB	-27,08 dB
Vokal 2	Bleeding 14,809-16,231	-38,15 dB	-61,57 dB	-61,65 dB	-61,67 dB
	Vokal aktif 11,479-12,901	-27,15 dB	-31,29 dB	-31,18 dB	-31,13 dB
Vokal 3	Bleeding 21,755-25,030	-30,81 dB	-59,80 dB	-60,27 dB	-60,56 dB
	Vokal aktif 25,003-28,278	-29,79 dB	-34,69 dB	-34,61 dB	-34,57 dB
Vokal 4	Bleeding 16,190-17,275	-40,28 dB	-56,18 dB	-55,99 dB	-55,87 dB
	Vokal aktif 17,274-18,360	-26,05 dB	-29,29 dB	-29,24 dB	-29,22 dB
Vokal 5	Bleeding 30,750-35,383	-34,33 dB	-57,53 dB	-57,64 dB	-57,69 dB
	Vokal aktif 26,117-30,750	-23,92 dB	-27,19 dB	-27,14 dB	-27,11 dB

Berdasarkan data pada ketiga tabel di atas, terlihat pola penurunan RMS yang konsisten seiring dengan bertambahnya intensitas pemrosesan. Pada saat parameter diatur di tingkat 50%, nilai RMS pada segmen *bleeding* (kebocoran suara latar) mengalami penurunan yang signifikan pada seluruh sampel. Hal ini menunjukkan efektivitas *plugin* dalam meredam suara gangguan latar panggung. Namun, peningkatan parameter pemrosesan 75% dan 100% ternyata turut memicu penurunan nilai RMS pada segmen vokal aktif. Penurunan ini rata-rata berkisar antara 2 hingga 3 dB jika dibandingkan dengan kondisi awal (0%). Menyusutnya nilai RMS vokal ini menjadi indikasi awal bahwa pemrosesan pada intensitas tinggi tidak hanya meredam gangguan latar, tetapi juga mulai mereduksi energi dari sinyal vokal utama. Untuk membuktikan rasio efektivitas antara reduksi gangguan dan bertahannya energi vokal tersebut, data RMS ini selanjutnya dikonversi menjadi nilai *Signal-to-Noise Ratio* (SNR).

B. *Hasil Signal-to-Noise Ratio (SNR)*

Setelah mendapatkan angka RMS pada tahap sebelumnya, langkah selanjutnya adalah menghitung nilai *Signal-to-Noise Ratio* (SNR).

Secara sederhana, perhitungan ini bertujuan untuk melihat seberapa jernih suara vokal penyanyi jika dibandingkan dengan suara bocoran instrumen di latar belakangnya. Semakin besar nilai SNR yang dihasilkan, maka suara vokalnya akan semakin bersih dan menonjol. Perhitungan ini dilakukan dalam mengurangi nilai RMS vokal aktif dengan nilai RMS *bleeding* (kebisingan).

Sebagai contoh, berikut adalah penjabaran cara menghitung nilai SNR pada sampel Vokal 1 Campursari berdasarkan data yang diambil dari Tabel 1:

1. Tingkat Reduksi 0%
   Diketahui:
   RMS *Signal* (vokal aktif) = -15,85 dB
   RMS *Noise* (*bleeding*) = -38,07 dB
   SNR = -15,85 –(-38,07) = 22,22 dB
2. Tingkat Reduksi 50%
   Diketahui:
   RMS *Signal* (vokal aktif) = -18,90 dB
   RMS *Noise* (*bleeding*) = -48,68 dB
   SNR = -18,90 –(-48,69) = 29,78 dB

Prosedur perhitungan yang sama diterapkan secara berulang di semua tingkat reduksi untuk seluruh sampel vokal dari ketiga *genre* musik. Hasil dari perhitungan *Signal-to-Noise Ratio* (SNR) secara menyeluruh disajikan pada Tabel 4, Tabel 5, dan Tabel 6.

Tabel 4. SNR Vokal Campursari

Sampel	SNR 0%	SNR 50%	SNR 75%	SNR 100%
Vokal 1	22,22 dB	29,78 dB	29,65 dB	29,57 dB
Vokal 2	13,35 dB	31,29 dB	31,35 dB	31,37 dB
Vokal 3	14,12 dB	43,42 dB	44,31 dB	44,86 dB
Vokal 4	7,31 dB	20,96 dB	20,92 dB	20,90 dB
Vokal 5	18,35 dB	26,22 dB	26,10 dB	26,03 dB

Tabel 5. SNR Vokal Dangdut

Sampel	SNR 0%	SNR 50%	SNR 75%	SNR 100%
Vokal 1	10,00 dB	19,30 dB	19,27 dB	19,27 dB
Vokal 2	18,96 dB	30,49 dB	30,57 dB	30,60 dB
Vokal 3	10,45 dB	45,70 dB	45,87 dB	46,18 dB
Vokal 4	15,56 dB	24,15 dB	24,05 dB	24,08 dB
Vokal 5	9,96 dB	22,31 dB	22,26 dB	22,22 dB

Tabel 6. SNR Vokal Karawitan

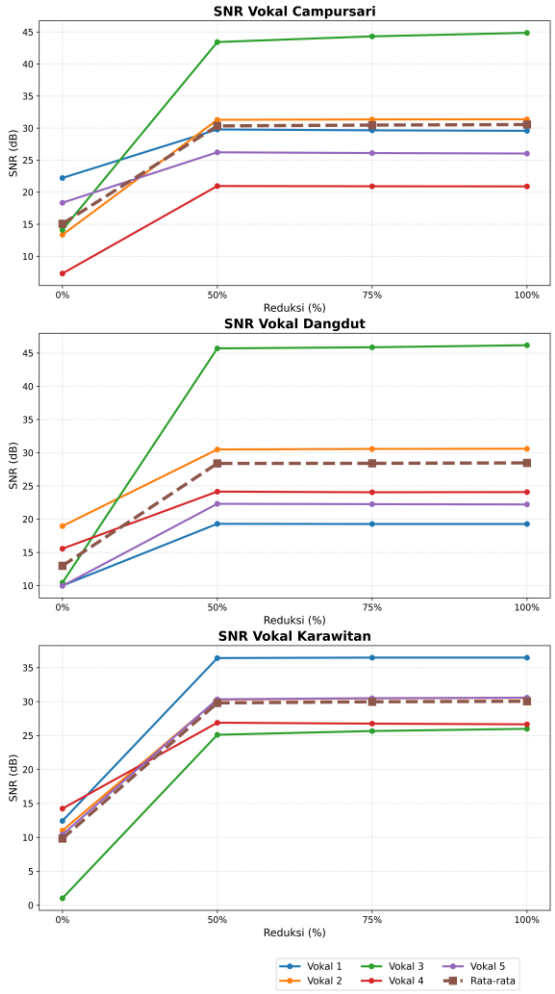
Sampel	SNR 0%	SNR 50%	SNR 75%	SNR 100%
Vokal 1	12,42 dB	36,41 dB	36,48 dB	36,48 dB
Vokal 2	11,00 dB	30,28 dB	30,47 dB	30,54 dB
Vokal 3	1,02 dB	25,11 dB	25,66 dB	25,99 dB
Vokal 4	14,23 dB	26,89 dB	26,75 dB	26,65 dB
Vokal 5	10,41 dB	30,34 dB	30,50 dB	30,58 dB

Berdasarkan hasil perhitungan pada ketiga tabel SNR di atas, penerapan *plugin* Clarity Vx pada intensitas 50% menghasilkan lonjakan nilai SNR yang tajam di seluruh sampel uji. Peningkatan ini membuktikan bahwa sinyal vokal yang awalnya tercampur dengan suara latar berhasil diisolasi dengan baik. Namun, peningkatan intensitas pemrosesan menuju 75% dan 100% tidak lagi diikuti oleh kenaikan nilai SNR yang sebanding. Pada rentang intensitas tersebut, pergerakan nilai SNR cenderung berhenti, bahkan hanya mengalami sedikit penurunan pada beberapa sampel. (2)

Fenomena ini sejalan dengan temuan penurunan nilai RMS vokal pada pengukuran sebelumnya. Ketika parameter pemrosesan ditekan hingga batas maksimal 100%, *plugin* tidak hanya menghilangkan kebisingan latar, tetapi juga ikut mereduksi kekuatan volume asli dari sinyal vokal. Akibatnya, rentang rasio kejernihan yang dihasilkan tidak lagi bisa bertambah. Hal ini mengindikasikan bahwa penggunaan parameter di tingkat 100% tidak memberikan tingkat efektivitas yang sepadan dan justru berisiko menurunkan integritas sinyal vokal aktif.

C. Evaluasi kinerja plugin Waves Clarity Vx

Evaluasi kinerja secara keseluruhan dilakukan dengan menganalisis nilai rata-rata SNR dari setiap *genre* musik. Penggabungan nilai ini bertujuan untuk memetakan pola umum kemampuan *plugin* Waves Clarity Vx dalam menangani berbagai karakteristik kebocoran suara panggung. Hasil rata-rata tersebut divisualisasikan pada grafik berikut.



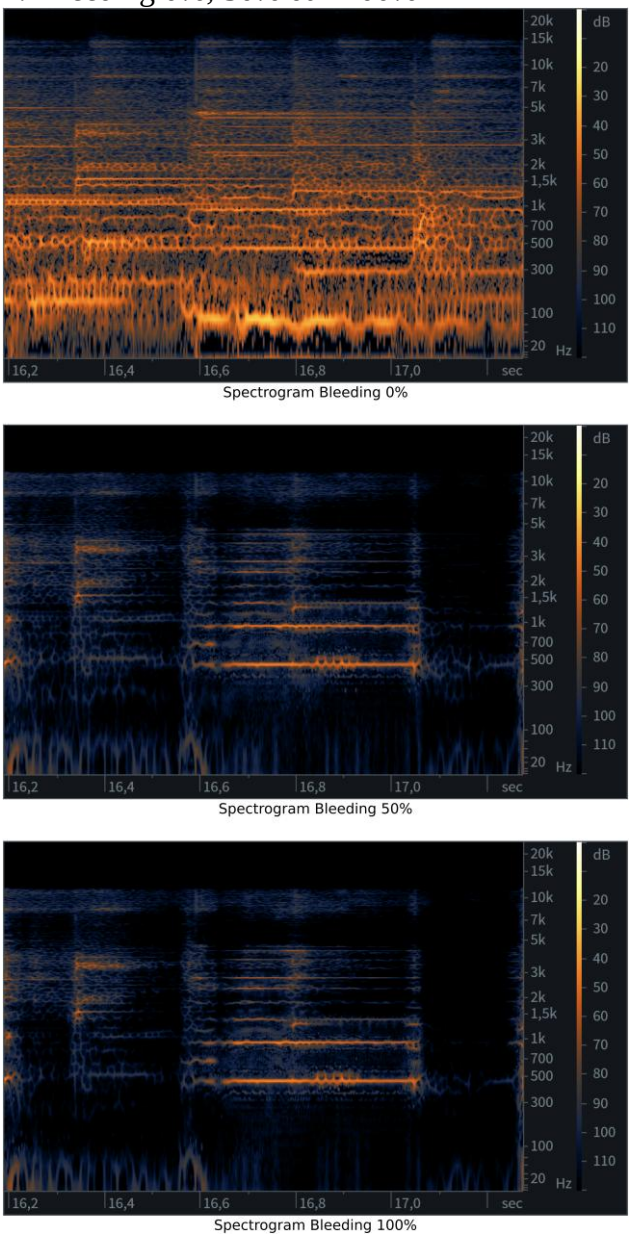
Gambar 5. Perbandingan SNR Berdasarkan Genre dan Tingkat Reduksi

Berdasarkan grafik tersebut, terdapat dua temuan utama dari pergerakan garis data. Pertama, grafik menunjukkan bahwa garis untuk *genre* Karawitan dan Dangdut memiliki sudut kenaikan yang paling tajam dari kondisi awal (0%) menuju intensitas 50%. Hal ini menandakan bahwa *plugin* memberikan margin peningkatan nilai SNR yang lebih besar pada kedua *genre* tersebut dibandingkan dengan genre Campursari yang arah garisnya naik secara lebih landai.

Temuan kedua terlihat dari arah pergerakan ketiga garis *genre* setelah melewati titik 50%. Grafik secara jelas memperlihatkan bahwa dari intensitas 50% menuju 75% hingga 100%, ketiga garis tersebut mengalami kelandaian dan bergerak mendatar. Pola *horizontal* pada grafik ini membuktikan bahwa peningkatan intensitas pemrosesan di atas 50% sudah tidak efektif lagi karena nilai kejernihan sinyal telah mencapai titik jenuh. Oleh karena itu, melalui tren grafik ini, pengaturan parameter di tingkat 50% adalah sebagai titik kerja paling optimal.

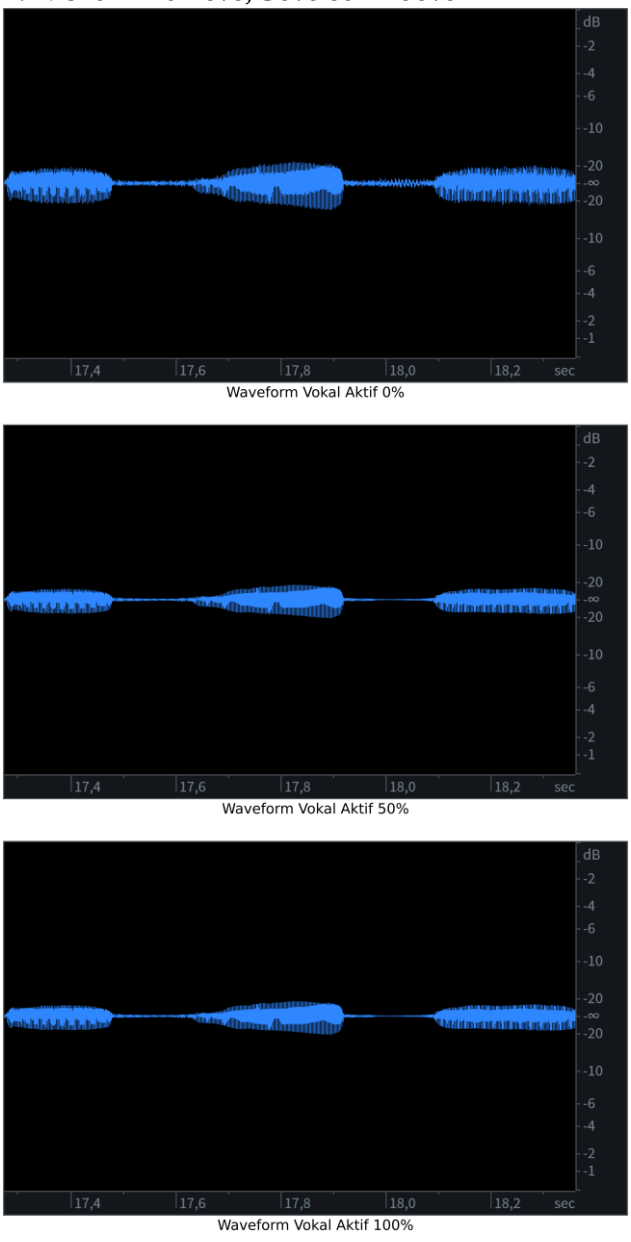
Untuk membuktikan alasan di balik mendatarnya garis grafik pada intensitas tinggi tersebut, evaluasi dilanjutkan dengan membandingkan visualisasi audio menggunakan perangkat lunak iZotope RX 11. Gambar berikut memperlihatkan perbandingan spektrum dari segmen kebocoran (*Bleeding*) dan bentuk gelombang (*Waveform*) dari segmen Vokal aktif pada tiga kondisi, yaitu 0%, 50%, dan 100%.

1. Bleeding 0%, 50% dan 100%



Gambar 6. Spectogram Bleeding

2. Vokal Aktif 0%, 50% dan 100%



Gambar 7. Waveform Vokal Aktif

Berdasarkan perbandingan visual tersebut, dampak dari intensitas pemrosesan dapat diamati secara objektif. Pada gambar spektrum *Bleeding* di kondisi awal (0%), area frekuensi terlihat sangat terang yang menandakan tingginya kebocoran suara alat musik. Saat diproses pada tingkat 50%, intensitas warna terang tersebut sebagian besar sudah berhasil dibersihkan. Kemudian pada tingkat 100%, layar spektrum menjadi jauh lebih redup. Meskipun masih terdapat sedikit warna samar yang merupakan sisa frekuensi (residu) akibat ketidaksempurnaan performa *plugin*, visualisasi ini membuktikan bahwa kebisingan latar berhasil diredam dengan baik.

Namun, temuan yang paling menarik terlihat pada bentuk gelombang (*Waveform*) Vokal Aktif. Pada kondisi awal (0%), gelombang suara vokal terlihat paling tebal karena masih bercampur dengan frekuensi kebocoran. Saat diproses ke 50%, ketebalan gelombang tersebut menyusut sebagai tanda bahwa *plugin* telah mengeliminasi suara latar pada vokal. Tetapi, ketika intensitas didorong maksimal hingga 100%, gelombang terlihat nyaris sama persis dan tidak menunjukkan perubahan visual yang berarti dibandingkan kondisi 50%. Kesamaan visual ini menjadi bukti akurat dengan data perhitungan, di mana nilai SNR dan RMS vokal terjadi perubahan yang sangat minim pada intensitas tinggi.

IV. KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa plugin AI Waves Clarity Vx memiliki efektivitas yang tinggi dalam mereduksi microphone bleeding pada kanal vokal live sound. Melalui pengujian kuantitatif, pengaturan parameter intensitas pada tingkat 50% terbukti menjadi titik kerja yang paling optimal. Pada kondisi tersebut, kebocoran suara instrumen latar berhasil diredam secara signifikan tanpa merusak komponen energi asli dari sinyal vokal utama. Sebaliknya, peningkatan parameter ke tingkat 75% hingga 100% menghasilkan tren data yang stagnan karena algoritma AI mulai mereduksi frekuensi vokal utama, yang dibuktikan oleh penurunan nilai RMS vokal serta bentuk gelombang (waveform) yang konstan.

Dari hasil pengujian simulasi Virtual Soundcheck ini, plugin Clarity Vx dinilai sangat layak dan efisien untuk diimplementasikan secara langsung pada sistem (*live sound*). Implikasi praktis dari temuan ini menunjukkan bahwa penggunaan teknologi pemisahan suara berbasis AI dapat menjadi solusi alternatif pengganti perangkat konvensional. Penerapan di lapangan sangat disarankan untuk mengunci

parameter intensitas pemrosesan pada batas aman 50% guna menjaga kualitas natural suara vokal serta menghindari risiko distorsi artikulasi. Untuk penelitian selanjutnya, disarankan melakukan pengujian langsung pada lingkungan panggung dengan variasi jenis mikrofon yang berbeda guna mengukur tingkat latensi pemrosesan dan stabilitas komputasi perangkat keras secara waktu nyata.

REFERENSI

[1] Y. Hashizume, L. Li, A. Miyashita, and T. Toda, "Learning Separated Representations for Instrument-based Music Similarity," *APSIPA Trans. Signal Inf. Process.*, vol. 14, no. 1, pp. 1–36, 2025, doi: 10.1561/116.20250013.

[2] G. Plaja-Roglans, X. Serra, and M. Rocamora, "Leveraging Carnatic Live Recordings for Singing Voice Separation Using Regression-Guided Latent Diffusion," *Proc. Int. Soc. Music Inf. Retr. Conf.*, vol. 2025, pp. 830–838, 2025, doi: 10.5281/zenodo.17706605.

[3] L. A. A. R. Putri, I. G. E. W. Pratama, I Dewa Made Bayu Atmaja Darmawan, A. A. I. N. Eka Karyawati, Ida Bagus Made Mahendra, and I Ketut Gede Suhartana, "Penerapan Metode Fast Independent Component Analysis (FastICA) dalam Memisahkan Vokal dan Instrumen Seni Geguntangan," *J. Buana Inform.*, vol. 13, no. 1, pp. 77–84, 2022, doi: 10.24002/jbi.v13i1.5693.

[4] R. Rajesh and P. Rajan, "Neural Networks for Interference Reduction in Multi-Track Recordings," *IEEE Work. Appl. Signal Process. to Audio Acoust.*, vol. 2023-Octob, pp. 1–5, 2023, doi: 10.1109/WASPAA58266.2023.10248133.

[5] Dewi Nurdiyah, Eko Mulyanto Yuniarno, Yoyon Kusnendar Suprapto, and Mauridhi Hery Purnomo, "IRAWNET: A Method for Transcribing Indonesian Classical Music Notes Directly from Multichannel Raw Audio," *Emit. Int. J. Eng. Technol.*, vol. 11, no. 2, pp. 246–264, 2023, doi: 10.24003/emitter.v11i2.827.

[6] R. A. Amin, "ANALISIS PERBANDINGAN KEBISINGAN TERHADAP SISTEM SUARA AKUSTIK DENGAN EFEK SURROUND DAN TANPA EFEK SURROUND," *J. Elektro Kontrol*, vol. 5, no. 1, pp. 36–48, Jul. 2025, doi: 10.24176/elkon.v5i1.14811.

[7] C. J. Steinmetz *et al.*, "Audio Signal Processing in the Artificial Intelligence Era: Challenges and Directions," *AES J. Audio Eng. Soc.*, vol. 73, no. 8, pp. 406–428, 2025, doi: 10.17743/jaes.2022.0209.

[8] E. In, R. Chandrakumar, F. Soria, and W. W. W. V. Com, "Lightweight Deep Neural Networks for Real-Time Speech Enhancement on Edge Devices Comparator With Low Offset Voltage Implemented in 45nm CMOS Technology," vol. 1, no. 3, pp. 1–8, 2025, doi: https://doi.org/10.17051/NJSAP/01.03.01.

[9] C. J. Steinmetz, J. Pons, S. Pascual, and J. Serrà, "Automatic multitrack mixing with a differentiable mixing console of neural audio effects," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2021-June, pp. 71–75, 2021, doi: 10.1109/ICASSP39728.2021.9414364.

[10] A. Défossez, "Hybrid Spectrogram and Waveform Source Separation," pp. 1–13, 2021. doi: 10.48550/arXiv.2111.03600

[11] S. Rouard, F. Massa, and A. Defossez, "Hybrid Transformers for Music Source Separation," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2023-June, pp. 2–6, 2023, doi: 10.1109/ICASSP49357.2023.10096956.

[12] H. Dubey *et al.*, "ICASSP 2023 Deep Noise Suppression Challenge," *IEEE Open J. Signal Process.*, vol. 5, pp. 725–737, 2024, doi: 10.1109/OJSP.2024.3378602.

[13] Waves.com," Clarity Vx Pro". https://www.waves.com/plugins/clarity-vx-pro#tab-product

[14] K. Zhen, M. S. Lee, J. Sung, S. Beack, and M. Kim, "Psychoacoustic Calibration of Loss Functions for Efficient End-to-End Neural Audio Coding," *IEEE Signal Process. Lett.*, vol. 27, pp. 2159–2163, 2020, doi: 10.1109/LSP.2020.3039765.