

Implementasi Metode ViSQOL Dalam Mengidentifikasi Noise pada Kualitas Suara Streaming Spotify

Jimmi Setiawan Matangkin^{*1}, Agustinus Rudatyo Himamunanto², Haeni Budiati³

^{1,2,3}Program Studi Informatika, Fakultas Sains dan Komputer, Universitas Kristen Immanuel Yogyakarta, Indonesia

Email: ¹setiawanjimmi23@gmail.com, ²rudatyo@ukrim.ac.id, ³heni@ukrimuniversity.ac.id

Abstrak

Kualitas suara pada layanan streaming Spotify seringkali tidak konsisten akibat gangguan noise dan variasi parameter jaringan, yang berdampak pada pengalaman pengguna (QoE). Penelitian ini bertujuan mengevaluasi kualitas suara Spotify menggunakan algoritma ViSQOL dengan menganalisis pengaruh noise dari pink noise, background noise, compression noise dan impulse noise, serta network noise yang ditentukan lewat parameter jaringan yaitu throughput, delay, packet loss dan jitter. Metode kuantitatif diterapkan dengan merekam 800 sampel audio menggunakan Audacity, kemudian dianalisis melalui ViSQOL di MATLAB untuk menghasilkan nilai MOS, SNR, dan Spectral Distortion, sementara parameter jaringan dipantau menggunakan Wireshark. Hasil menunjukkan pink noise 50% menurunkan MOS hingga 61-65%, impulse noise berdampak paling buruk pada MOS di angka 15-17%, dan background noise relatif dapat ditoleransi dengan nilai MOS yaitu 61-65%. Delay berpengaruh lebih signifikan terhadap QoE dibanding packet loss, dengan MOS mencapai 3.78 pada delay 62.15 ms. Temuan ini memberikan wawasan bagi pengembang layanan streaming untuk mengoptimalkan mitigasi noise dan manajemen jaringan, serta menjadi referensi bagi penelitian terkait evaluasi kualitas audio berbasis persepsi pengguna.

Kata kunci: Kualitas Suara, Mean Opinion Score, Parameter Jaringan, ViSQOL.

Implementation of the ViSQOL Method in Identifying Noise in Spotify Streaming Sound Quality

Abstract

The sound quality on Spotify streaming service is often inconsistent due to noise interference and network parameter variations, which impacts user experience (QoE). This study aims to evaluate Spotify sound quality using the ViSQOL algorithm by analyzing the effect of noise from pink noise, background noise, compression noise and impulse noise, as well as network noise determined through network parameters, namely throughput, delay, packet loss and jitter. The quantitative method was applied by recording 800 audio samples using Audacity, then analyzed through ViSQOL in MATLAB to produce MOS, SNR, and Spectral Distortion values, while network parameters were monitored using Wireshark. The results showed that 50% pink noise reduced MOS by 61-65%, impulse noise had the worst impact on MOS at 15-17%, and background noise was relatively tolerable with a MOS value of 61-65%. Delay had a more significant effect on QoE than packet loss, with MOS reaching 3.78 at a delay of 62.15 ms. These findings provide insights for streaming service developers to optimize noise mitigation and network management, as well as serve as a reference for research related to user perception-based audio quality evaluation.

Keywords: Mean Opinion Score, Network Parameters, ViSQOL, Voice Quality.

1. PENDAHULUAN

Kemajuan teknologi digital dalam industri musik telah mengubah masyarakat dalam mengakses dan menikmati berbagai genre musik. Layanan streaming musik seperti Spotify menjadi pilihan utama, di kalangan generasi muda. Spotify adalah salah satu platform musik, yang menawarkan akses ke jutaan lagu dengan berbagai pilihan kualitas suara, mulai dari low yaitu 24 kbps hingga very high dengan 320 kbps, bagi pengguna premium[1]. Meskipun Spotify sudah memberikan banyak kemudahan untuk mendengarkan musik, tetapi pengguna masih sering melaporkan ketidakpuasan mereka terhadap kualitas suara yang didengarkan. Perubahan kualitas ini sering dipengaruhi oleh beberapa faktor teknis seperti kecepatan internet, jenis perangkat, dan pengaturan audio, sehingga menimbulkan tanggapan mengenai Quality of Experience (QoE) pengguna[2]. QoE merupakan penilaian persepsi, preferensi, dan kepuasan pengguna terhadap layanan yang diberikan[3]. Perbedaan - Perbedaan kualitas ini tidak selalu dapat dirasakan secara signifikan oleh pengguna,

bahkan beberapa fitur seperti normalisasi volume justru dapat mengurangi rentang dinamis suara sehingga kejernihan dan kenyamanan mendengarkan menjadi berkurang. Selain itu, kualitas suara juga sering kali tidak konsisten antar perangkat atau kondisi jaringan. Misalnya, pada koneksi internet yang lambat, Spotify akan secara otomatis menurunkan bitrate audio untuk menjaga kelancaran streaming, yang berdampak pada penurunan kualitas suara. Hal ini menjadi permasalahan utama bagi pengguna yang mengutamakan kualitas audio dalam menikmati musik. Permasalahan ini belum banyak diangkat secara mendalam dalam penelitian-penelitian sebelumnya, terutama yang berfokus pada pengalaman pengguna Spotify di Indonesia. Karena kebanyakan peneliti hanya berfokus pada platform lain seperti YouTube Music, maka dari itu penelitian ini mengevaluasi kualitas suara Spotify, di kondisi jaringan dan dengan MOS untuk melihat sejauh mana penilaian preferensi pengguna yang mungkin berbeda-beda.

Pendekatan yang digunakan untuk mengevaluasi kualitas suara secara objektif salah satunya dengan algoritma ViSQOL (Virtual Speech Quality Objective Listener). Algoritme ViSQOL dilatih agar dapat mensimulasikan persepsi pendengaran manusia, sehingga mampu memberikan penilaian yang mendekati pengalaman mendengarkan musik bagi pengguna[4]. Algoritma ini bekerja dengan menirukan proses biologis manusia dalam mendengarkan suara, lalu prosesnya adalah dengan membandingkan sinyal audio asli dengan sinyal yang terdegradasi, dan menghasilkan skor kualitas berdasarkan kemiripan persepsi manusia. Namun dalam penelitian ini evaluasi ViSQOL dilakukan tanpa adanya sinyal referensi (audio asli), penelitian yang menunjukkan bahwa pendekatan non-intrusive ViSQOL tetap dapat memberikan hasil yang konsisten dengan penilaian manusia[5]. Dengan demikian, ViSQOL dapat menjadi alat untuk mengevaluasi konsistensi kualitas suara Spotify, dalam berbagai kondisi teknis, seperti perubahan bitrate akibat jaringan atau perbedaan perangkat.

Kajian tentang evaluasi kualitas audio berbasis persepsi manusia menunjukkan perkembangan yang sangat signifikan dalam metrik objektif, seperti penelitian yang berjudul "ViSQOL v3: An Open Source Objective Speech and Audio Metric", peneliti memperkenalkan inovasi ViSQOL v3 sebagai alat evaluasi audio berbasis neurogram similarity (NSIM) dan MOS dengan fitur presisi teknis seperti time alignment, di mana kedua studi ini menegaskan pentingnya pendekatan berbasis persepsi manusia[6]. Penelitian ini sejalan dengan penelitian yang berjudul "Analisis Kepuasan Pengguna pada Layanan Musik Streaming Aplikasi YouTube Music Mobile Menggunakan Model EUCS", peneliti mengungkap bahwa aspek teknis seperti system speed dan accuracy, yang terkait dengan presisi evaluasi objektif, serta elemen subjektif seperti ease of use, berpengaruh terhadap kepuasan pengguna, sehingga dapat mengintegrasikan metrik teknis, psikometrik, dan analisis pengguna untuk pengembangan layanan multimedia yang optimal, sekaligus menekankan perlunya pendekatan multidisiplin dalam riset evaluasi kualitas audio[7].

Meskipun penelitian-penelitian sebelumnya, yang telah berhasil mengembangkan metode evaluasi objektif berbasis persepsi manusia, serta telah mengintegrasikan aspek teknis dan subjektif dalam analisis kepuasan pengguna, tetapi terdapat beberapa perbedaan mendasar dengan penelitian ini. Penelitian yang telah mereka lakukan juga masih berfokus pada evaluasi kualitas audio dalam kondisi laboratorium atau jaringan stabil, maka penelitian ini dilakukan dengan mengevaluasi kualitas audio Spotify dalam konteks noise yang dapat mempengaruhi kualitas suara. Noise tersebut dibagi menjadi 4 jenis yaitu Background Noise yang sering disebut Suara latar yang mengganggu, serta Impulse Noise yaitu Gangguan tiba-tiba yang biasa terjadi seperti adanya notifikasi[8]. Lalu ada Compression Noise atau gangguan akibat kompresi audio, dan Network Noise yang biasa disebut gangguan karena paket data yang hilang saat streaming. Penelitian sebelumnya juga cenderung memisahkan antara analisis objektif seperti ViSQOL dengan pengukuran kepuasan pengguna, maka penelitian ini mengusulkan algoritme ViSQOL yang dapat menghubungkan secara langsung hasil pengukuran objektif dengan persepsi subjektif pengguna. Selain itu, belum banyak penelitian yang secara nyata menggunakan algoritme ViSQOL untuk mengevaluasi kualitas suara streaming Spotify pada berbagai kondisi perangkat dan jaringan. Padahal, kualitas suara yang konsisten dan memuaskan sangat penting dalam membangun kepuasan dan loyalitas pengguna di tengah persaingan layanan streaming musik yang semakin ketat. Dengan demikian, penelitian yang menganalisis kualitas suara Spotify secara objektif dan subjektif menjadi sangat relevan untuk dilakukan. Penelitian terdahulu yang telah dijabarkan, cenderung memisahkan analisis objektif dan pengukuran kepuasan pengguna, sementara penelitian ini mengusulkan pendekatan langsung antara keduanya untuk menilai konsistensi kualitas suara Spotify. Berdasarkan tinjauan literatur yang telah didapatkan ada beberapa gap yang menjadi fokus dalam penelitian ini. Pertama dalam penerapan ViSQOL untuk evaluasi kualitas suara Spotify. Kedua penelitian sebelumnya tidak ada yang menerapkan hasil dari evaluasi ViSQOL dengan parameter jaringan untuk memprediksi QoE pengguna. Ketiga beberapa peneliti belum Menerapkan simulasi noise dengan 3 jenis noise sekaligus untuk menilai dan mengevaluasi kualitas audio.

Penelitian ini bertujuan untuk menganalisis kualitas suara streaming Spotify yang telah direkam melalui aplikasi Audacity, dan menggunakan algoritme ViSQOL untuk mengevaluasi hasil rekaman audio tersebut, dengan mempertimbangkan 3 jenis noise yang akan dimasukkan ke dalam hasil rekaman audio, serta berbagai faktor dalam parameter jaringan yang dapat mempengaruhi konsistensi dan kepuasan pengguna dalam menikmati layanan musik digital. Penelitian ini dapat memberikan kontribusi dalam pengembangan layanan streaming musik yang lebih baik, serta menjadi referensi bagi penelitian-penelitian selanjutnya di bidang kualitas layanan audio digital.

2. METODE PENELITIAN

2.1. Desain Penelitian

Penelitian ini menggunakan metode pemrosesan yaitu analisis kualitas suara. Evaluasi terhadap kualitas suara merupakan alat untuk menganalisis secara kuantitatif berdasarkan perasaan manusia[9]. Pemrosesan dilakukan dengan menggunakan perangkat lunak MATLAB, yang memungkinkan evaluasi suara dan menganalisis kualitas suara. Serta dilakukan pengumpulan dan peninjauan berbagai literatur yang relevan dengan kualitas suara pada layanan streaming multimedia, khususnya Spotify. Metode kuantitatif adalah sebuah metode penelitian yang berfokus pada analisis numerik dari kumpulan data[10]. Dalam penelitian ini, metode kuantitatif digunakan untuk pengumpulan data numerik dari hasil penggunaan algoritme ViSQOL dengan didapatkan nilai MOS, SNR, dan Spectral Distortion. Penilaian akhir didapatkan dengan menghitung rata-rata dari hasil algoritme ViSQOL yang digunakan untuk menganalisis kualitas suara. Untuk menghitung rata-rata tersebut, digunakanlah Persamaan (1) :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (1)$$

Persamaan (1) merupakan perhitungan untuk mencari rata-rata dari hasil ViSQOL. $\bar{X}$ merupakan hasil rata-rata nilai MOS, SNR dan Spectral Distortion. X_i adalah Nilai metrik ke- i dari ViSQOL. n merupakan Jumlah sampel (800 rekaman).

2.2. Populasi dan Sample Data

penelitian ini, terdapat populasi data dan sample data. Populasi data dalam penelitian ini adalah kualitas suara yang merambat melalui sinyal suara pada jaringan atau perangkat yang digunakan selama streaming. Kualitas suara tersebut meliputi suara streaming yang direkam pada aplikasi Spotify. Sampel data yang dipergunakan dalam penelitian ini yaitu rekaman sinyal suara streaming pada aplikasi Spotify dan pengambilan parameter jaringan seperti Throughput, Delay, Packet Loss, dan Jitter untuk dilihat pengaruhnya pada kualitas suara selama streaming berlangsung. Proses pengambilan sampel dilakukan dengan kondisi jaringan yang stabil dan kondisi jaringan yang buruk melalui jenis jaringan yaitu Wi-Fi. Setiap sample audio hasil rekaman memiliki durasi 1 menit di setiap rekaman, dengan format audio yaitu .Wav yang berjumlah 800 rekaman, dan berjumlah 200 di setiap rentang waktu pengambilan rekaman.

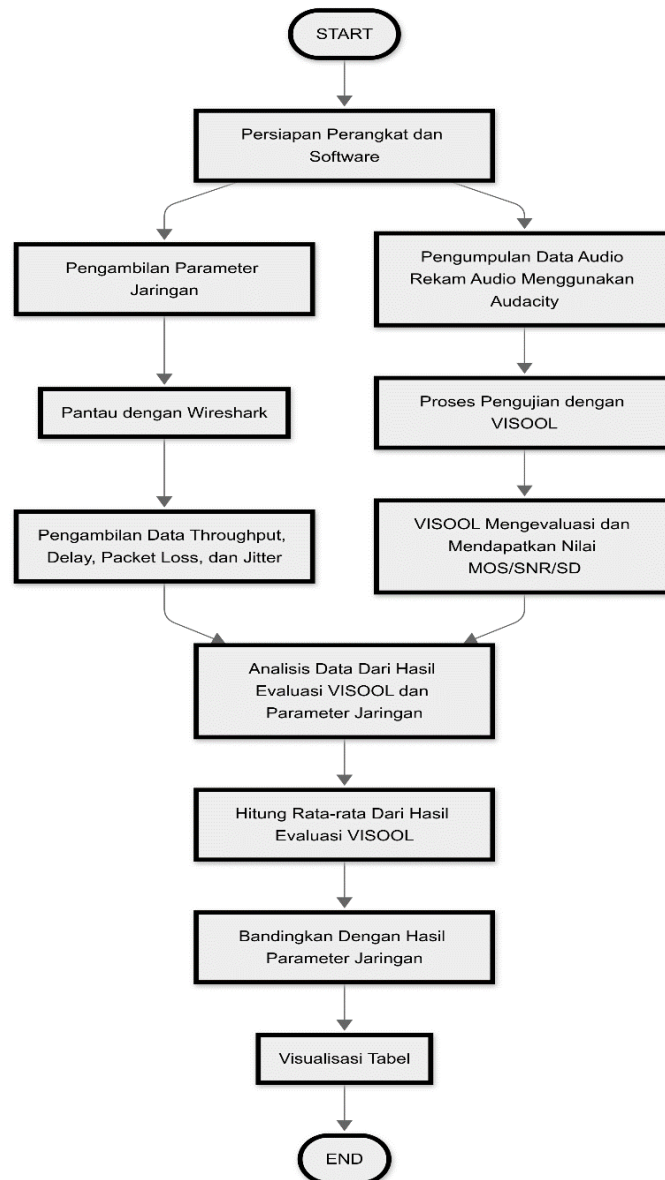
2.3. Spesifikasi Kebutuhan

- A. Perangkat Keras: PC dengan spesifikasi prosesor Intel(R) Core(TM) i3-4160 CPU @ 3.60GHz, RAM 4 GB.
- B. Perangkat Lunak:
 - Audacity versi terbaru digunakan untuk memutar dan merekam streaming audio Spotify.
 - MATLAB R2023a digunakan untuk mengolah serta mengevaluasi data audio dengan menggunakan algoritme ViSQOL.
- C. Jaringan: Pengujian dilakukan menggunakan jaringan Wi-Fi. Dengan aplikasi Wireshark versi 4.4.6 untuk mensimulasikan pengaruh kualitas jaringan terhadap kualitas suara streaming Spotify.

2.4. Prosedur pengujian

Prosedur pengujian dirancang berdasarkan alur yang melibatkan beberapa tahapan, pertama yaitu pengumpulan data rekaman dilakukan saat streaming Spotify, lalu direkam menggunakan Audacity dalam kondisi waktu tertentu seperti pada jam produktif, jam istirahat, dan pada saat malam hari, serta pengumpulan data parameter jaringan seperti Throughput, Delay, Packet Loss, dan Jitter yang diambil dari aplikasi Wireshark. Data audio yang digunakan meliputi suara dari streaming Spotify. Kedua pelaksanaan pengujian, setelah didapatkan rekaman suara dari streaming dan data dari parameter jaringan. Lalu dilakukan pengujian pada data rekaman audio dengan menggunakan algoritme ViSQOL untuk melihat nilai MOS, nilai Signal to Noise Ratio (SNR) dan nilai Spectral Distortion (SD), dengan tingkat intensitas pink noise 10% dan 50% di setiap pengujian. Pink Noise merupakan salah satu stimulus untuk membandingkan pengaruh jenis noise terhadap persepsi identitas suara[11]. Dalam pengujian ini jenis – jenis noise yaitu compression noise, background noise dan impulse noise digunakan untuk mensimulasikan gangguan yang sering terjadi di lingkungan pengguna. Serta network noise yang adalah parameter jaringan seperti Throughput, Delay, Packet Loss, dan Jitter diuji berdasarkan hasil pengamatan menggunakan aplikasi Wireshark. Ketiga yaitu Pengumpulan hasil pengujian, data rekaman audio yang telah didapatkan melalui evaluasi algoritme ViSQOL, serta parameter – parameter jaringan yang didapatkan dari aplikasi wireshark, lalu dianalisis lebih mendalam untuk mendapatkan nilai rata – rata MOS, SNR dan Spectral Distortion, serta Untuk parameter jaringan setelah nilai Throughput, Delay, Packet Loss dan Jitter didapatkan maka dilakukan penilaian menggunakan E-Model untuk mendapatkan nilai MOS yang sempurna. Penelitian ini

dirancang, pada alur kerja pengujian kualitas suara dalam layanan streaming Spotify. Sistem ini bertujuan untuk mengukur dan menganalisis kualitas suara yang diterima oleh pengguna melalui metode objektif ViSQOL. Sesuai dengan prosedur yang telah dijelaskan, maka dibuatlah flowchart, yang dapat dilihat pada Gambar 1



Gambar 1. Prosedur Pengujian

2.5. Proses Normalisasi Data

Nilai MOS, SNR dan Spectral Distortion untuk mempermudah memahami data dari hasil pengukuran, seluruh metrik kualitas audio dinormalisasi ke skala persentase (0-100%) dengan pendekatan sebagai berikut:

1. Mean Opinion Score (MOS) Dikonversi dari skala 1-5 ke persentase 0-100% menggunakan formula:

$$MOS_{\%} = \frac{(MOS_{raw}-1)}{4} \times 100 \quad (2)$$

Terlihat pada Persamaan (2) terdapat MOS_{raw} yaitu Nilai MOS asli dengan skala 1 hingga 5 dan $MOS_{\%}$ adalah nilai MOS yang ternormalisasi.

2. Signal-to-Noise Ratio (SNR) dinormalisasi dari satuan dB dengan asumsi nilai maksimum yaitu 40 dB.

$$SNR_{\%} = \frac{SNR_{dB}}{SNR_{max}} \times 100 \quad (3)$$

Pada Persamaan (3), SNR_{dB} adalah Nilai SNR dalam desibel (dB) yang dibagi dengan nilai skala SNR_{max} yang sudah ditetapkan. SNR_{max} adalah Nilai skala yang telah ditetapkan yaitu 40 dB, hasilnya untuk normalisasi ke persentase dan $SNR_{\%}$ merupakan nilai SNR yang telah ternormalisasi.

3. Spectral Distortion (SD) untuk menormalisasikan nilainya yang dari desibel (dB) menjadi persen, maka diperlukan skala dengan batas maksimum 20 dB.

$$SD_{\%} = 1 - \frac{SD_{dB}}{SD_{max}} \times 100 \quad (4)$$

Normalisasi untuk nilai Spectral Distortion (SD) menggunakan Persamaan (4). Pada Persamaan (4), SD_{dB} merupakan nilai distorsi spektral dalam skala desibel (dB), SD_{Max} adalah batas distorsi untuk menghitung skala menjadi persen dengan skala maksimum yaitu 20dB, dengan begitu nilai $SD_{\%}$ merupakan nilai SD ternormalisasi yang dinyatakan dalam persen.

2.6. Teknik Analisis Data

Setelah penilaian kualitas suara dikumpulkan, data dianalisis untuk dilihat, penggunaan algoritme ViSQOL dalam mengevaluasi kualitas suara yang diterima oleh pengguna Spotify dapat dinilai secara efektif, serta data penilaian parameter jaringan digunakan untuk melihat faktor-faktor yang mempengaruhi kualitas pada aplikasi Spotify. Implementasi ViSQOL dalam penelitian ini digunakan untuk menilai pengalaman pengguna saat menggunakan layanan dan mendengarkan streaming musik. Untuk mendapatkan penilaian pengalaman pengguna dalam mendengarkan streaming musik digunakanlah algoritme ViSQOL pada MATLAB. Dengan ViSQOL maka didapatkan nilai SNR dan Spectral Distortion, serta nilai MOS untuk dianalisis lebih lanjut. Lalu dengan aplikasi wireshark digunakan untuk melihat parameter jaringan seperti Throughput, Delay, Packet Loss dan Jitter saat streaming Spotify.

3. HASIL PENGUJIAN

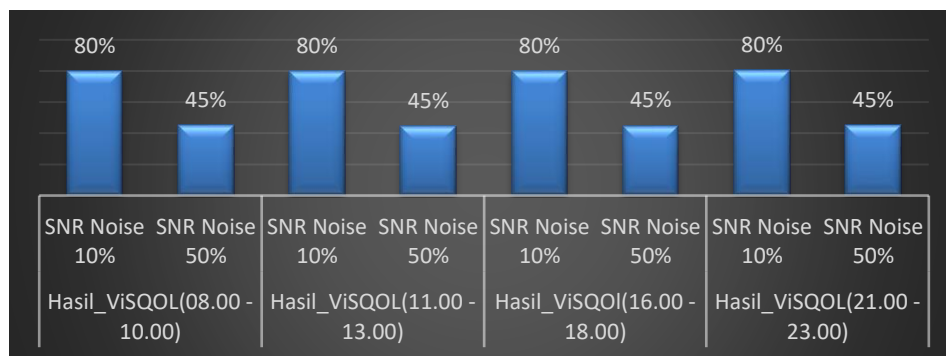
Penelitian ini dilakukan untuk mengevaluasi dan menilai kualitas suara streaming Spotify yang telah direkam dahulu. Lalu dievaluasi oleh algoritme ViSQOL untuk menghasilkan nilai MOS, SNR dan Spectral Distortion. Serta mensimulasikan pengaruh dari setiap gangguan noise terhadap sinyal suara yang telah direkam. Gangguan jaringan juga diujikan dengan penilaian pada parameter jaringan, lalu hasil parameter jaringan tersebut dianalisis lebih lanjut untuk dilihat nilai MOS yang merupakan penilaian persepsi pengguna. Hasil-hasil dari evaluasi, serta penilaian terhadap gangguan noise dan penilaian pada parameter-parameter jaringan yang telah didapatkan, bisa dilihat pada setiap subbab yang memberikan petunjuk gambar dan tabel.

3.1. Hasil Evaluasi ViSQOL Dan Hasil Pengujian Jenis-Jenis Noise



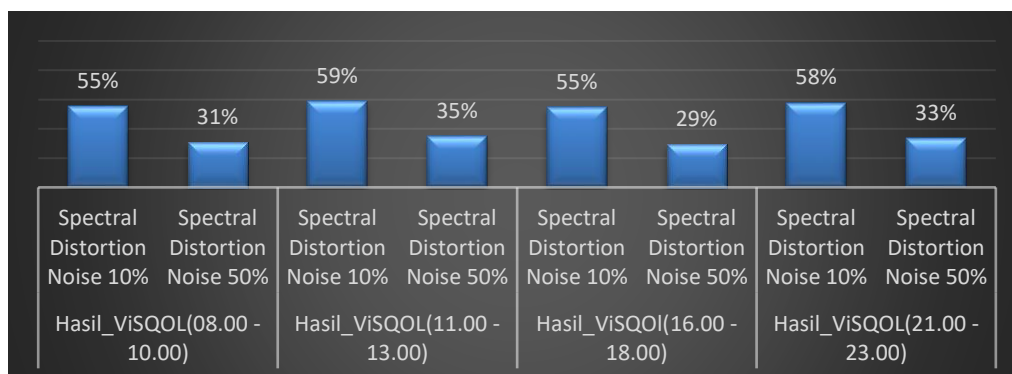
Gambar 2. Hasil Pengujian ViSQOL Nilai MOS (%)

Pengujian menggunakan algoritma ViSQOL, pada Gambar 2 menunjukkan bahwa tingkat gangguan pink noise memiliki dampak signifikan terhadap persepsi kualitas suara yang diukur melalui nilai MOS. Dalam simulasi ini, pink noise disisipkan ke dalam sinyal audio pada tingkat 10% dan 50%, yang kemudian dievaluasi oleh ViSQOL untuk menilai degradasi kualitas suara secara objektif. Hasilnya menunjukkan bahwa pada level pink noise 10%, nilai MOS berada pada kisaran 80–83%, menandakan kualitas suara yang masih sangat baik dan dapat diterima oleh pendengar. Sebaliknya, ketika pink noise ditingkatkan menjadi 50%, nilai MOS turun drastis menjadi 61–65%, yang mengindikasikan penurunan signifikan dalam kenyamanan mendengarkan.



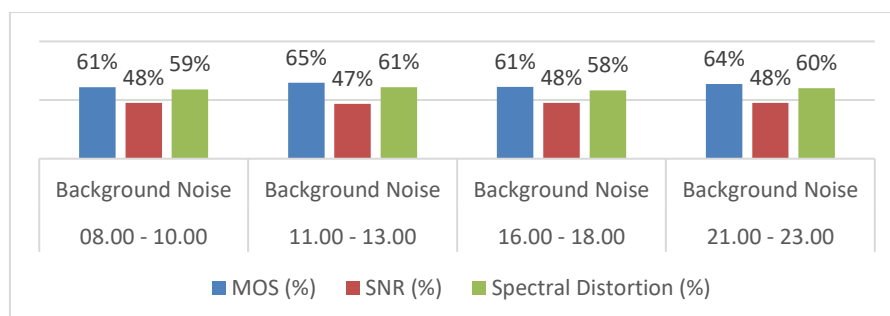
Gambar 3. Hasil Pengujian ViSQOL Nilai SNR (%)

Hasil evaluasi ViSQOL terhadap kualitas sinyal berdasarkan rasio SNR secara konsisten menunjukkan bahwa tingkat gangguan yang lebih tinggi secara drastis menurunkan kejernihan sinyal. Saat pink noise disimulasikan pada level 10%, nilai SNR tetap stabil di angka 80%, dengan angka tersebut sinyal audio yang diberikan gangguan pink noise dapat dinyatakan bisa mempertahankan kualitas audio dengan baik. Namun, saat level noise meningkat menjadi 50%, nilai SNR dinyatakan turun hingga 45% secara merata di seluruh waktu pengujian, ini menandakan bahwa peningkatan level noise dapat menurunkan nilai sinyal audio yang cukup signifikan.



Gambar 4. Hasil Pengujian ViSQOL Spectral Distortion (%)

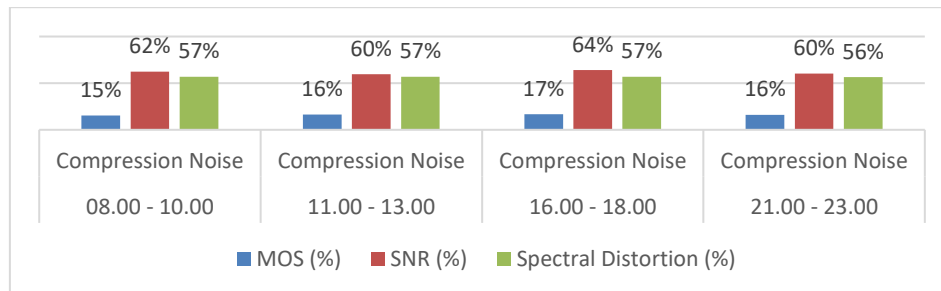
Gambar 4 merupakan hasil pengujian Spectral Distortion menggunakan algoritma ViSQOL gambar tersebut menunjukkan hasil yang beragam, terlihat pada nilai distorsi spektral justru lebih tinggi saat pink noise berada di tingkat 10% dibandingkan 50%. Hal ini menandakan bahwa pada level noise rendah, perbedaan spektral antara sinyal asli dan hasil gangguan noise lebih menonjol terdeteksi oleh algoritma, mungkin sinyal masih mendominasi namun mengalami perubahan halus yang cukup signifikan. Sebaliknya, saat noise mencapai 50%, sinyal cenderung tenggelam dalam kebisingan, sehingga perbedaan spektral menjadi kurang terdeteksi secara tajam. Hasil ini menyatakan bahwa ViSQOL, dapat mendeteksi detail spektral halus dalam kondisi sinyal yang relatif bersih, dan sekaligus menegaskan pentingnya mempertimbangkan konteks sinyal dengan noise saat mengevaluasi metrik distorsi spektral.



Gambar 5. Hasil Pengujian Background Noise (%)

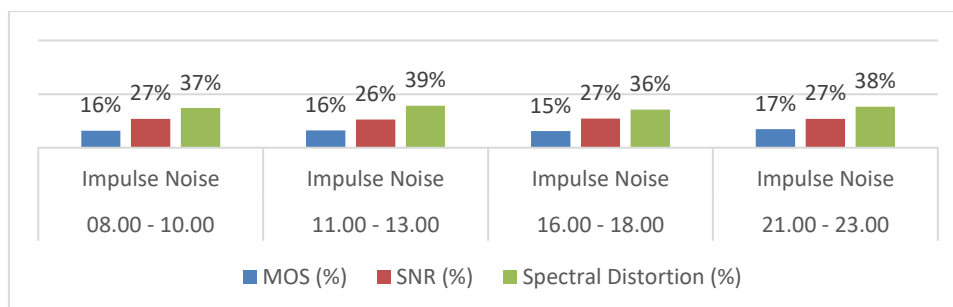
Hasil pengujian terhadap gangguan Background Noise, pada Gambar 5 menunjukkan bahwa meskipun terjadi gangguan dari suara latar seperti suara angin, keramaian, atau suara lingkungan lainnya, kualitas suara masih berada pada

tingkat yang relatif stabil dan dapat diterima oleh pengguna[12]. Hal ini bisa dilihat dari nilai MOS yang konsisten tinggi di angka 61% hingga 65% pada seluruh rentang waktu pengujian, ini menunjukkan bahwa persepsi pengguna terhadap suara masih baik. Meskipun nilai SNR berada pada kisaran rendah di angka 47% dan 48%, yang menunjukkan adanya gangguan terhadap kejernihan sinyal, dampaknya terhadap persepsi pengguna tidak terlalu signifikan. Serta nilai Spectral Distortion yang cukup tinggi yaitu 58% dan 61%, hasil ini memperlihatkan adanya perubahan spektral suara akibat gangguan latar, namun tidak menyebabkan degradasi yang dapat menurunkan pengalaman pengguna.



Gambar 6. Hasil Pengujian Compression Noise (%)

Gambar 6 merupakan hasil pengujian pada gangguan Compression Noise, dalam konteks layanan streaming, kompresi bertujuan untuk memperkecil ukuran file agar mudah ditransmisikan melalui jaringan, namun hal ini sering kali mengorbankan kualitas suara[13]. Terlihat dalam nilai Spectral Distortion di kisaran angka 56% dan 57%, sementara SNR cenderung berada pada nilai yang tinggi, berkisar antara 60% sampai 64%, yang menandakan kualitas sinyal suara yang tetap kuat meskipun ada noise. Namun MOS yang merupakan penilaian persepsi pengguna mendapatkan nilai yang cukup rendah yaitu 15% hingga 17%, hal ini menandakan bahwa persepsi pengguna terhadap kualitas suara tetap buruk, terlepas dari parameter teknis lainnya.



Gambar 7. Hasil Pengujian Impulse Noise (%)

Impulse noise merupakan gangguan suara yang bersifat tiba-tiba dan tajam, seperti bunyi klik, dentuman, atau letupan pendek, yang cenderung sangat mengganggu meskipun hanya terjadi dalam waktu singkat[14]. Berdasarkan Gambar 7, terlihat bahwa nilai Mean Opinion Score (MOS) yang sangat rendah berkisar antara 15% di sore hari dan naik hingga 17% di malam hari, ini menunjukkan nilai persepsi pengguna terhadap kualitas suara yang sangat terganggu. Nilai sinyal suara juga dapat dilihat pada nilai SNR yang rendah dengan angka 26% hingga 27% dan diikuti oleh nilai Spectral Distortion yang cukup tinggi di angka 36% hingga 39%, ini menandakan bahwa Impulse noise sangat mengacaukan kejernihan sinyal suara, serta gangguan tersebut menyebabkan nilai distorsi tinggi pada struktur frekuensi suara asli.

3.2. Hasil Analisis Parameter Jaringan

Tabel 1. Parameter Jaringan

Waktu	Throughput	Packet Loss	Avg. Delay (ms)	Avg. Jitter (ms)
08.00 – 10.00	99,18 k	0,60%	131,67 ms	0,00357 ms
11.00 – 13.00	199 k	1,90%	62,15 ms	0,00360 ms
16.00 – 18.00	62 k	0,56%	175,24 ms	0,00360 ms
21.00 – 23.00	94,95 k	0,49%	132,16 ms	-0,00626 ms

Tabel 1 merupakan hasil pengamatan dari aplikasi Wireshark. Wireshark merupakan sebuah perangkat lunak analisis jaringan berbasis open-source, yang memungkinkan pengguna untuk memantau, merekam, dan menganalisis data yang

dikirimkan melalui jaringan komputer[15]. Berdasarkan hasil analisis yang ditampilkan Tabel I, terlihat bahwa gangguan jaringan yang biasa disebut network noise memiliki nilai yang bervariasi sepanjang hari. Nilai packet loss tertinggi tercatat pada pukul 11.00–13.00 sebesar 1.9%, meskipun throughput-nya tinggi, mengindikasikan bahwa volume data besar tidak menjamin kestabilan transmisi. Sementara itu, delay tertinggi terjadi pada pukul 16.00–18.00, mencapai 175 ms, yang berpotensi menyebabkan keterlambatan suara saat streaming. Pada Tabel 1 nilai jitter terdeteksi sangat kecil dan konsisten di angka 0.0036 ms, sehingga tidak menjadi kontributor utama penurunan kualitas suara. Hasil ini menegaskan bahwa dalam konteks noise jaringan, faktor packet loss dan delay lebih berdampak nyata terhadap degradasi kualitas audio, khususnya saat layanan streaming digunakan dalam kondisi jaringan yang padat atau tidak stabil[16].

3.3. Pembahasan Hasil Pengujian

Berdasarkan hasil pengujian yang telah dilakukan, dapat disimpulkan bahwa berbagai jenis noise dan berdasarkan parameter jaringan yang mempengaruhi kualitas suara streaming Spotify, serta dengan tingkat dampak yang berbeda-beda. Pink noise menunjukkan pengaruh yang signifikan terhadap penilaian kualitas suara. Pada tingkat gangguan 10%, nilai MOS masih berada dalam kisaran 80% sampai 83%, yang mengindikasikan kualitas suara yang baik dan dapat diterima oleh pengguna. Namun, ketika noise ditingkatkan menjadi 50%, nilai MOS turun drastis menjadi 61% hingga 65%, yang menunjukkan bahwa penurunan persentase ini berdampak pada kualitas audio. Hasil ini sejalan dengan penelitian yang menyatakan bahwa gangguan noise dapat mengurangi kejernihan audio dan memengaruhi persepsi pendengaran manusia[17]. Selain itu, nilai SNR yang turun dari 80% menjadi 45% pada tingkat noise yang lebih tinggi memperkuat temuan bahwa gangguan noise secara signifikan mengurangi kualitas sinyal audio.

Analisis dari ketiga kategori noise mengungkap dampak berbeda pada kualitas audio. Dari Gambar 5 kita bisa melihat dampak dari Background noise, meskipun dampaknya pada nilai SNR yang mengalami penurunan diangka 47% disiang hari hingga 48% dimalam hari, angka presentase ini menunjukkan adanya gangguan terhadap kejernihan sinyal suara, tetapi masih bisa ditoleransi, dan peningkatan Spectral Distortion diangka presentase 58% sampai 61% yang dapat dinyatakan tidak terlalu memengaruhi persepsi pengguna dengan MOS tetap stabil di 61% hingga 65%, ini menunjukkan bahwa gangguan suara latar relatif dapat ditoleransi. Hasil temuan dari pengujian Background noise, dapat disimpulkan sejalan dengan penelitian, yang menemukan bahwa sebagian besar penilai MOS terhadap kualitas suara yang telah melalui proses reduksi noise tetap terdengar jelas dan bersih, meskipun masih terdapat gangguan ringan pada beberapa jenis noise[18]. Sebaliknya, compression noise pada Gambar 6 terlihat menghasilkan nilai presentase SNR yang cukup baik dengan kisaran diangka 60% dan 64%, namun nilai persepsi manusia yaitu MOS cenderung sangat turun diangka 15% hingga 17%, yang menandakan bahwa kompresi audio dapat menurunkan penilaian persepsi pengguna. Pada Gambar 7 terlihat juga nilai MOS sangat turun diangka 15% dan 17%, yang disebabkan oleh impulse noise. Serta nilai sinyal suara yang terdegradasi oleh impulse noise menyebabkan nilai SNR juga turun dengan angka persentase hanya di 26% hingga 27%, ini membuktikan bahwa gangguan tiba-tiba seperti klik atau dentuman sangat mengganggu kenyamanan pengguna[19].

Tabel 2. Parameter MOS

Jam	8.00 – 10.00	11.00 – 13.00	16.00 – 18.00	21.00 – 23.00
Nilai Parameter MOS	1,31	3,78	3,86	3,91

Pengukuran parameter jaringan yang terlihat di Tabel 1 dan nilai Parameter MOS pada Tabel 2 mengungkap dampak network noise terhadap kualitas audio streaming. Hasil pengukuran menunjukkan bahwa pada periode 08.00-10.00, meskipun throughput mencapai 99.18 kbps, kombinasi packet loss 0.6% dengan delay 131.67 ms menghasilkan MOS 1.31 yang termasuk dalam kategori tidak memuaskan. Pada rentang waktu, diperiode 11.00-13.00 dengan packet loss lebih tinggi (1.9%) namun delay lebih rendah (62.15 ms) justru menghasilkan MOS 3.78 yang secara signifikan lebih baik, menunjukkan bahwa delay merupakan faktor yang lebih berpengaruh dalam pembentukan network noise dibandingkan packet loss. Terlihat pada Tabel 2, Stabilitasnya nilai MOS pada malam hari diangka 3.91 dan meski throughput rendah diangka 94.95 kbps, tetapi dapat membuktikan bahwa minimnya network noise pada packet loss 0.49%, serta delay 132.16 ms, ini lebih krusial daripada bandwidth tinggi. Sementara itu terlihat ada keterbatasan dalam penelitian ini yang terletak pada lingkup pengujian pada jaringan WiFi, sehingga memerlukan kajian lebih lanjut pada berbagai jenis jaringan dan kondisi yang lebih bervariasi[20].

4. KESIMPULAN

Penelitian ini menyatakan bahwa metode ViSQOL dapat secara efektif mengukur kualitas audio Spotify yang diakibatkan oleh noise dan variasi jaringan. Hasil menunjukkan bahwa pink noise dan impulse noise paling berdampak pada penurunan MOS, sementara background noise masih dapat diterima, oleh telinga pendengar, dengan nilai MOS yang tinggi. Analisis terhadap hasil parameter jaringan mengungkap bahwa delay lebih berpengaruh pada kualitas audio

dibandingkan dengan packet loss, dalam penilaian pengalaman pengguna. Temuan ini memberikan wawasan baru tentang perbedaan antara metrik objektif yaitu SNR dan Spectral Distortion dengan penilaian persepsi subjektif yaitu MOS, yang sering diabaikan dalam evaluasi kualitas audio berbasis jaringan.

Hasil pengujian menunjukkan bahwa Pink noise 50% lebih signifikan menurunkan MOS hingga diangka 61-65%, dibanding 10% diangka 80-83%. Impulse noise berdampak paling buruk pada nilai MOS dikisaran 15-17% meskipun terjadi secara tidak teratur. Delay yang lebih kecil nilainya dibawah 150 ms mendapatkan nilai MOS hingga 1.31, sementara packet loss 1.9% dengan delay 62.15 ms tetap mempertahankan MOS diangka 3.78. Temuan ini juga mengungkap ketidak stabilnya antara metrik teknis yaitu SNR yang diangka 45% pada pink noise 50% dan persepsi manusia yaitu MOS yang terprediksi diangka 61-65%.

Implikasi praktis penelitian ini menekankan perlunya pengembangan algoritma adaptif yang memprioritaskan mitigasi impulse noise dan optimasi manajemen delay. Untuk penelitian selanjutnya, perlu dilakukan eksperimen komparatif pada berbagai platform streaming (YouTube Music, Apple Music) dengan sampel pengguna yang lebih beragam. Studi lanjutan juga dapat mengintegrasikan machine learning untuk memprediksi MOS berbasis parameter jaringan secara real-time.

DAFTAR PUSTAKA

- [1] G. Azzahira, P. Zakaria, T. Pradekso, and L. R. Rahmijai, "the Influence of Price Perception , Advertising Perception , Audio Quality Perception , Music Download Perception , and Unlimited Skip Perception on the Decision To Use Spotify Premium Application," *J. Ilm. Univ. Diponegoro*, 2023, [Online]. Available: <https://ejournal3.undip.ac.id/index.php/interaksi-online/article/view/45288/31465>
- [2] M. R. dos Santos, A. P. Batista, R. L. Rosa, M. Saadi, D. C. Melgarejo, and D. Z. Rodríguez, "AsQM: Audio Streaming Quality Metric Based on Network Impairments and User Preferences," *IEEE Trans. Consum. Electron.*, vol. 69, no. 3, pp. 408–420, 2023, doi: 10.1109/TCE.2023.3255411.
- [3] F. P. E. Putra, M. Aziz, G. Arifin, A. Rohman, A. Rizki, and A. M. Syam, "Analisis Kualitas Jaringan Integratif Quality of Service (QoS) dan Quality of Experience (QoE) pada Layanan Multimedia di jaringan 5G," *semanTIK*, vol. 10, no. 1, p. 175, 2024, doi: 10.55679/semanTIK.v10i1.46554.
- [4] L. F. Schill, T. Piechowiak, C. Laroche, and P. Mowlae, "Comparing neural network architectures for non-intrusive speech quality prediction," *Speech Commun.*, vol. 165, 2024, doi: 10.1016/j.specom.2024.103123.
- [5] A. Kasperuk and S. K. Zieliński, "Non-intrusive method for audio quality assessment of lossy-compressed music recordings using convolutional neural networks," *Int. J. Electron. Telecommun.*, vol. 70, no. 2, pp. 331–339, 2024, doi: 10.24425/ijet.2024.149549.
- [6] M. Chinen, F. S. C. Lim, J. Skoglund, N. Gureev, F. O’Gorman, and A. Hines, "ViSQOL v3: An Open Source Production Ready Objective Speech and Audio Metric," *2020 12th Int. Conf. Qual. Multimed. Exp. QoMEX 2020*, 2020, doi: 10.1109/QoMEX48832.2020.9123150.
- [7] A. Chaerunnisa, "Analisis Kepuasan Pengguna pada Layanan Musik Streaming Aplikasi YouTube Music Mobile Menggunakan Model EUCS," *Institutional Repos. UIN Syarif Hidayatullah*, pp. 1–23, 2024.
- [8] O. Barkovska and A. Havrashenko, "Kharkiv National University of Radio Electronics , Kharkiv , Ukraine RESEARCH OF THE IMPACT OF NOISE REDUCTION METHODS," pp. 57–65, 2024, doi: 10.18664/ikszt.v29i3.313606.
- [9] I. Salsabila, S. Anwar, and R. Radhiah, "Perbandingan Kualitas Suara Smartphone Menggunakan Metode Dynamic Time Warping (DTW)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 5, no. 1, pp. 82–90, 2021, doi: 10.29207/resti.v5i1.2764.
- [10] Sofwatillah, Risnita, M. S. Jailani, and D. A. Saksitha, "Teknik Analisis Data Kuantitatif dan Kualitatif dalam Penelitian Ilmiah," *J. Genta Mulia*, vol. 15, no. 2, pp. 79–91, 2024.
- [11] D. Murwaningrum and R. A. Indrayuana, "Pink Noise Sebagai Guide Dalam Mixing," *Sorai J. Pengkaj. dan Pencipta. Musik*, vol. 16, no. 1, pp. 45–54, 2023, doi: 10.33153/sorai.v16i1.5086.
- [12] C. Ick and B. McFee, "Sound event detection in urban audio with single and multi-rate PCEN," *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.*, vol. 2021-June, pp. 880–884, 2021, doi: 10.1109/ICASSP39728.2021.9414697.
- [13] J. Nistal, S. Lattner, and G. Richard, "Conditioning Using Generative Adversarial Networks," 2021.
- [14] S. Sadrizadeh, N. Zarmehi, E. A. Kangarshahi, H. Abin, and F. Marvasti, "A Fast Iterative Method for Removing Impulsive Noise from Sparse Signals," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 1, pp. 38–48, 2021, doi: 10.1109/TCSVT.2020.2969563.

- [15] R. Soepeno, "Wireshark: An Effective Tool for Network Analysis," *CYBV - Introd. Methods Netw. Anal.*, no. September, pp. 1–15, 2023, doi: 10.13140/RG.2.2.34444.69769.
- [16] K. Masykuroh, A. D. Ramadhani, and N. Iryani, "Analisis Qos Dan Qoe Pada Video Pembelajaran Online Di Institut Teknologi Telkom Purwokerto (Ittp)," *Transmisi*, vol. 23, no. 2, pp. 40–47, 2021, doi: 10.14710/transmisi.23.2.40-47.
- [17] Y. Zhou, Y. Liu, and H. Niu, "Perceptual Characteristics of Voice Identification in Noisy Environments," *Appl. Sci.*, vol. 12, no. 23, 2022, doi: 10.3390/app122312129.
- [18] A. Nogales, J. Caracuel-Cayuela, and Á. J. García-Tejedor, "Analyzing the Influence of Diverse Background Noises on Voice Transmission: A Deep Learning Approach to Noise Suppression," *Appl. Sci.*, vol. 14, no. 2, 2024, doi: 10.3390/app14020740.
- [19] A. Ragano, "Data-driven Quality of Experience for Digital Audio Archives," 2022.
- [20] A. Panca and R. Rustam, "Analisis Performansi Codec G.711 Dan G.729 Berbasis Issabel Menggunakan Metode MOS E-Model," *JOINS (Journal Inf. Syst.*, vol. 8, no. 2, pp. 115–124, 2023, doi: 10.33633/joins.v8i2.8534.