



Analisis Kualitas Audio Steganografi MP3 Menggunakan Teknik Masking Pada Spectrogram

MP3 Steganography Audio Quality Analysis Using Spectrogram Masking Technique

Permadi Kusuma¹, Yudi Prayudi²

^{1,2}Program Studi Informatika, Fakultas Teknologi Industri, Universitas Islam Indonesia, Yogyakarta, Indonesia

¹23917007@students.uii.ac.id, ²prayudi@uii.ac.id

Abstract

Steganography hides secret information through media such as audio, video and images. This research aims to apply masking techniques to the audio spectrogram of MP3 media and analyze the results of steganography using the Python programming approach with the Librosa library. The problem in this research is how to insert a secret message into an audio file without disturbing the sound quality, and the secret message can be detected using masking techniques. So that the insertion process is carried out by utilizing certain frequency areas in the spectrogram which tend to be insensitive to human perception. The audio samples used in this study consist of the song Bernadya - Life Must Keep Going, ringtones Samsung Galaxy S3 and Samsung Galaxy S20. Audio analysis is done through spectrogram visualization, Mel-Frequency Cepstral Coefficients (MFCC) feature extraction, Zero-Crossing Rate (ZCR), and audio file metadata evaluation. The purpose of this analysis is to see if there are any strange or unnatural patterns that appear due to the insertion of secret messages in audio files. So, when there is data inserted in a hidden way, it is normal for small differences to appear in the form of lines, noise, or unnatural changes in visual structure. The results show that message insertion is successfully performed using masking techniques, but it causes changes in the visual and statistical characteristics of the original audio signal with the result of steganography. Meanwhile, the sound quality is not much different from the original audio.

Keywords: *Audio Steganography; Spectrogram; Masking; Spectrogram Analysis; Librosa*

Abstrak

Steganografi menyembunyikan informasi rahasia melalui media seperti audio, video dan gambar. Penelitian ini bertujuan untuk menerapkan teknik masking pada *spectrogram* audio dari media MP3, serta menganalisis hasil steganografi menggunakan pendekatan pemrograman *Python* dengan pustaka *Librosa*. Permasalahan pada penelitian ini ialah bagaimana menyisipkan pesan rahasia ke dalam file audio tanpa mengganggu kualitas suara, dan pesan rahasia dapat dideteksi menggunakan teknik masking. Sehingga proses penyisipan dilakukan dengan memanfaatkan area frekuensi tertentu pada spectrogram yang cenderung tidak sensitif terhadap persepsi manusia. Sampel audio yang digunakan dalam penelitian ini terdiri dari lagu *Bernadya - Hidup Harus Tetap Berjalan*, nada dering *Samsung Galaxy S3* dan *Samsung Galaxy S20*. Analisis audio dilakukan melalui visualisasi spectrogram, ekstraksi fitur *Mel-Frequency Cepstral Coefficients (MFCC)*, *Zero-Crossing Rate (ZCR)*, serta evaluasi metadata file audio. Tujuan dari analisis ini adalah untuk melihat apakah ada pola-pola aneh atau tidak wajar yang muncul akibat penyisipan pesan rahasia dalam file audio. Jadi ketika ada data disisipkan secara tersembunyi, biasanya akan muncul perbedaan kecil dalam bentuk garis, noise, atau perubahan struktur visual yang tidak alami. Hasil penelitian menunjukkan bahwa penyisipan pesan berhasil dilakukan menggunakan teknik masking, namun menimbulkan perubahan pada karakteristik visual dan statistik sinyal audio asli dengan hasil steganografi. Sedangkan kualitas suara tidak jauh berbeda dengan audio asli.

Kata kunci: Steganografi Audio; Spectrogram; Masking; Spectrogram Analysis; Librosa

1. Pendahuluan

Steganografi merupakan ilmu untuk menyembunyikan pesan ke dalam media lain agar orang selain pengirim dan penerima tidak mengetahui keberadaan pesan

tersebut [1]. Steganografi pada audio berperan penting dalam keamanan informasi seperti digunakan untuk menyembunyikan data rahasia perusahaan, data pribadi, atau informasi sensitif lainnya. Namun sering kali ditemukan kasus kejahatan yang memanfaatkan

Steganografi. Menurut situs *cyberthreat.id*, kejahatan steganografi audio pernah terjadi pada tahun 2019 yaitu perusahaan keamanan siber *Symantec* dan *BlackBerry Cylance* melaporkan bahwa kelompok peretas menggunakan file audio *WAV* untuk menyembunyikan kode berbahaya. Pada kasus ini, malware disisipkan ke dalam file audio menggunakan teknik steganografi [2]. Dalam konteks audio forensik, teknik *Masking* pada steganografi audio dipilih karena menawarkan beberapa keuntungan, seperti pesan steganografi yang ada pada audio aman dari modifikasi dan bisa menampung pesan dengan jumlah kata atau kalimat yang bisa disesuaikan dengan kebutuhan. Penelitian tentang steganografi audio menggunakan teknik masking memberikan kontribusi terhadap *Forensik Digital (FD)*, terutama dalam deteksi dan analisis pesan tersembunyi. Teknik ini memungkinkan pesan disisipkan dalam audio dengan cara memasukkan pesan ke area spectrogram yang dimana untuk menampilkan area tersebut membutuhkan *software* khusus yang tidak semua orang tahu. Selain itu, teknik masking ini bisa meningkatkan keamanan komunikasi tersembunyi yang tahan dari modifikasi seperti kompresi dan mengubah format audio.

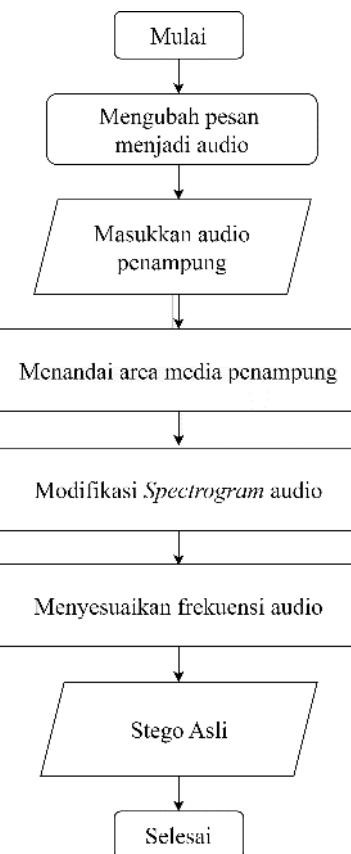
Steganografi audio yang ada perangkat lunak dapat menyematkan pesan dalam *WAV*, *AU*, dan bahkan suara *MP3* file. Menanamkan pesan rahasia dalam suara digital biasanya proses yang lebih sulit daripada menyematkan pesan di media lain, seperti gambar digital [3]. Kelebihan steganografi daripada kriptografi adalah pesan-pesannya tidak menarik perhatian orang lain, karena pesan-pesan tersebut dimasukan ke sebuah media penampung, sedangkan kriptografi hanya mengenkripsi pesan tersebut yang bisa saja menimbulkan kecurigaan orang lain [4]. Ada dua proses steganografi, yaitu embedding dan extracting. Proses embedding adalah proses penyisipan pesan rahasia ke dalam cover image. Pada proses ini dibutuhkan dua input, yaitu cover image dan secret message. Penyisipan secret message ke dalam cover image menghasilkan stego image. Proses extracting adalah proses pengambilan pesan rahasia dari stego image. Pada proses ini dibutuhkan satu input, yaitu stego image. Setelah proses extracting selesai akan dihasilkan secret message [5].

Salah satu media steganografi digital yang sering digunakan adalah audio. Audio steganografi adalah seni menyembunyikan pesan tersembunyi di dalam file audio dengan cara mengeksplorasi kelemahan indra pendengaran manusia [6]. Adapun alasan ilmiah kenapa penulis tertarik untuk meneliti steganografi audio terutama menggunakan sampel musik adalah musik memiliki spektrum frekuensi yang kompleks dengan berbagai nada, harmoni, dan dinamika. Selain itu, musik adalah media yang sering dipertukarkan di internet, membuatnya ideal untuk komunikasi rahasia tanpa menimbulkan kecurigaan. Sehingga proses

penyisipan pesan rahasia pada lagu dan nada dering dapat dikirim sebagai file biasa tanpa menarik perhatian.

2. Metode Penelitian

Pada penelitian ini, proses penyisipan dilakukan melalui area *Spectrogram* audio dari lagu dan nada dering berformat mp3 melalui teknik masking. Proses penyisipan tersebut dapat dilihat pada *Flowchart* gambar 1 di bawah ini.



Gambar 1. Flowchart Teknik Masking

Metodologi masking diadaptasi dari penelitian [7], dan [8], yang mengacu pada pendekatan secara umum untuk mendeteksi pesan rahasia. Penelitian ini menggunakan teknik masking karena teknik ini memanfaatkan kelemahan pendengaran manusia yaitu, suara lemah bisa “tertutupi” (dimasking) oleh suara yang lebih kuat di sekitar frekuensinya. Jadi, pesan bisa disisipkan tanpa terdengar, tapi tetap ada dalam data audio. Terdapat metode lainnya yaitu DCT (Discrete Cosine Transform) dan LSB (Least Significant Byte). Pada penelitian [7], menunjukkan teknik masking lebih kuat dalam pengujian ketahanan pesan seperti memotong bagian audio, melakukan kompresi dan mengubah format audio. Penelitian ini bertujuan untuk mengimplementasikan teknik masking dalam steganografi audio untuk menyisipkan pesan tersembunyi ke dalam spectrogram, membandingkan karakteristik hasil audio steganografi dengan audio asli



melalui visualisasi spectrogram, parameter akustik, dan metadata file. Mengetahui sejauh mana teknik masking pada steganografi audio dapat digunakan dalam konteks investigasi forensik digital, khususnya dalam proses deteksi atau identifikasi pesan tersembunyi pada audio.

Teknik masking berkembang dari konsep yang didasarkan pada *psikoakustik* dan teori sistem pendengaran manusia (*Human Auditory System, HAS*). Teknik masking dalam *psikoakustik* pertama kali diperkenalkan oleh *Hermann von Helmholtz*, seorang ilmuwan Jerman abad ke-19. Dalam steganografi audio, masking digunakan untuk menyembunyikan informasi rahasia dalam sinyal suara tanpa mengubah kualitas audio yang dapat didengar manusia.

Menurut [9] analisis *Spectrogram* ini didasarkan pada perbandingan pola yang dihasilkan dari kata-kata yang diucapkan oleh suara sampel dengan kata-kata yang diucapkan oleh suara barang bukti. *Spectrogram* membentuk pola-pola yang khas, sehingga dapat dijadikan sebagai salah satu analisis untuk audio forensik. Pada penyisipan pesan dilakukan dengan menyisipkan pesan rahasia ke dalam *Spectrogram* audio. Kemudian membandingkan antara audio sesudah dan sebelum disisipi dengan informasi rahasia. Dengan begitu bisa mengetahui kualitas audio setelah disisipi dengan informasi yang sudah diubah dalam format audio.

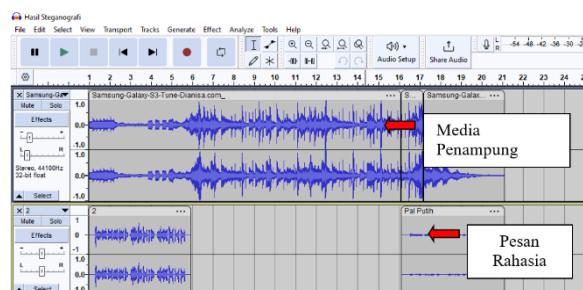
2.1. Mengubah Pesan Menjadi Audio

Steganografi yang umum digunakan adalah menyembunyikan informasi pada media gambar, di mana informasi text dimasukkan ke dalam bit pixel gambar. Namun metode yang sering digunakan masih cukup sederhana sehingga pihak ketiga masih bisa mendapatkan informasi yang disembunyikan [10]. Untuk menghindari permasalahan tersebut maka terciptalah steganografi yaitu metode menyembunyikan informasi pada sebuah media, bisa berupa media gambar, suara ataupun video. Pada penelitian ini menggunakan audio dari lagu “*Bernadya - Untungnya, Hidup Harus Tetap Berjalan*”, “Nada Dering Samsung Galaxy S3 dan nada dering *Samsung Galaxy S20*” yang ketiga file audio tersebut berformat mp3 yang nanti akan disisipkan pesan rahasia dalam bentuk kalimat dan gambar. Selanjutnya pesan rahasia dibuat menggunakan aplikasi *CoagulaLight1666*. Pesan stego awalnya berformat bitmap yang kemudian diubah menjadi audio agar bisa disisipkan ke area spectrogram menggunakan aplikasi *Audacity*. Aplikasi *Audacity* dapat digunakan sebagai perekam audio, memanipulasi serta mengedit suara. Selain itu dalam aplikasi *Audacity* ini memiliki banyak fitur menarik sehingga dapat meningkatkan ketangkapan dan pemahaman akan teknologi yang sedang berkembang [11]. Kelebihan dari aplikasi *Audacity* adalah fitur dan kestabilan. *Library* yang digunakan juga tidak terlalu

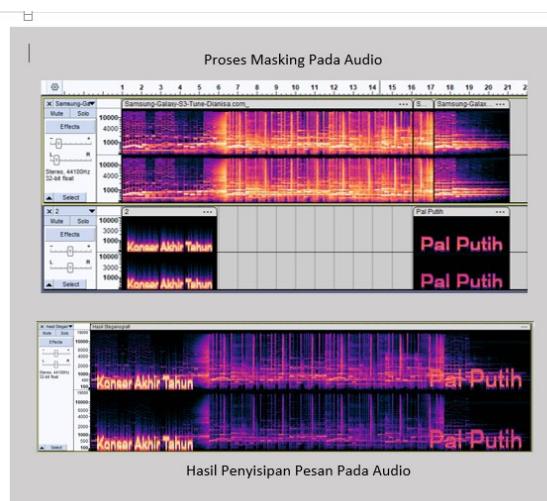
banyak dan waktu tunggunya yang tidak terlalu lama [12].

2.2. Penyisipan Pesan

Pesan rahasia tidak diubah menjadi bentuk karakter unik seperti halnya kriptografi. Pesan tersebut hanya disembunyikan ke dalam suatu media berupa musik, gambar, teks, atau media tampung digital lainnya dan terlihat seperti pesan atau berkas biasa [13]. Pada penelitian ini, peneliti menentukan bagian dari spektrum audio di mana suara tambahan dapat dimasukkan tanpa terdeteksi oleh pendengaran manusia. Biasanya, suara dengan frekuensi yang lebih tinggi atau yang berada di belakang suara yang lebih keras dipilih untuk penyisipan. Terdapat tiga pesan yang akan disisipkan pada tiga file audio. Pesan yang pertama bertuliskan “*Konser Akhir Tahun, Pal Putih, dan 31 Desember*” yang disisipkan pada audio dari lagu *Bernadya - Untungnya, Hidup Harus Tetap Berjalan*. Pesan kedua bertuliskan “*nama saya permadi kusuma konsentrasi forensika digital universitas islam indonesia*” yang disisipkan pada nada dering *Samsung Galaxy S3*. Untuk pesan yang ketiga dalam bentuk gambar yaitu gambar *Flashdisk* merk *Sandisk* yang disisipkan pada nada dering *Samsung Galaxy S20*. Proses penyisipan pesan stego menggunakan aplikasi *Audacity*.



Gambar 2. Proses penyisipan Steganografi pada Audio



Gambar 3. Tampilan Proses Steganografi Audio di Audacity



Selanjutnya pesan rahasia kemudian disisipkan ke dalam file audio pada bagian yang sudah dipilih dan difilter sebelumnya. Teknik penyisipan bisa menggunakan berbagai teknik, seperti menyisipkan pesan ke dalam area *Spectrogram* audio sebagai area yang ingin di masking. Dengan memasukkan nada pada frekuensi tingkat daya yang rendah, penyisipan data tersembunyi yang diekstraksi tercapai [14].

2.3. Penyimpanan dan Analisis

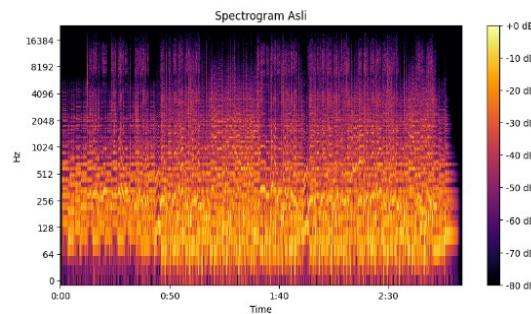
Dalam penelitian ini, dilakukan analisis *MFCC* (*Mel-Frequency Cepstral Coefficients*), *ZCR* (*Zero-Crossing Rate*), dan analisis *Spectrogram*. Ketiga analisis tersebut menggunakan pemrograman *Python* melalui *Google Colab*. Sedangkan analisis metadata file menggunakan *Software Mediainfo*. Pada penelitian [15] melakukan analisis audio stego menggunakan metode *LSB* dengan menganalisis area spektrogram. Sedangkan pada penelitian [16] menggunakan analisis metadata untuk mengidentifikasi tanda-tanda manipulasi pada file audio yang mengandung steganografi.

3. Hasil dan Pembahasan

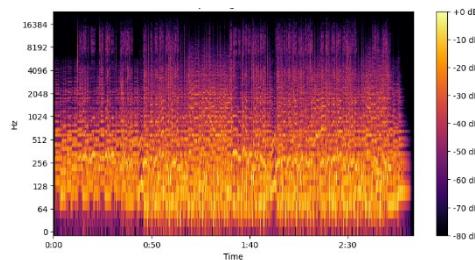
Pada penelitian ini akan mengimplementasikan steganografi audio dengan menyisipkan pesan ke dalam area *Spectrogram* audio. Penerapan steganografi pada file audio dianggap lebih rumit dibandingkan pada file video. Hal ini dikarenakan kemampuan pendengaran manusia (*Human Auditory System*) lebih peka daripada kemampuan pengelihatan manusia (*Human Visual System*) [17]. Untuk mendukung hasil implementasi steganografi audio pada *Spectrogram*, maka dilakukan analisis karakteristik audio yang secara umum sudah dilakukan oleh penelitian sebelumnya seperti, analisis perbedaan *Spectrogram*, *MFCC* (*Mel-Frequency Cepstral Coefficients*), *Zero-Crossing Rate* (*ZCR*) dan analisis metadata. Tujuan dari analisis ini adalah untuk melihat apakah ada pola-pola aneh atau tidak wajar yang muncul akibat penyisipan pesan rahasia dalam file audio. *Spectrogram* memvisualisasikan bagaimana energi suara tersebar dalam frekuensi dan waktu, jadi ketika ada data disisipkan secara tersembunyi, biasanya akan muncul perbedaan kecil dalam bentuk garis, noise, atau perubahan struktur visual yang tidak alami.

Alasan penulis memakai pendekatan berlapis yaitu *spectrogram*, *MFCC*, *ZCR*, dan metadata dikarenakan masing-masing memberikan sudut pandang berbeda dalam menganalisis perubahan pada audio. *Spectrogram* membantu melihat perubahan visual akibat penyisipan pesan, *MFCC* menangkap perbedaan karakteristik suara berdasarkan persepsi pendengaran manusia, *ZCR* mendeteksi adanya gangguan atau noise halus pada sinyal, dan metadata digunakan untuk memeriksa apakah ada perubahan teknis pada file. Dengan menggabungkan semua analisis ini, hasil

evaluasi menjadi lebih lengkap, akurat, dan dapat dipercaya.



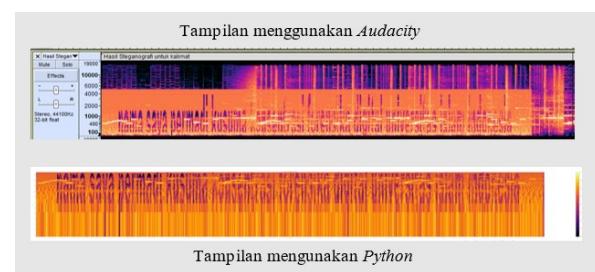
Gambar 4. *Spectrogram* audio asli



Gambar 5. *Spectrogram* audio stego

3.1. Analisis Steganografi Perbedaan Spectrogram

Pada tahapan ini analisis *Spectrogram* dilakukan guna melihat pola gelombang suara yang diucapkan dari formant setiap kalimat. Pada tampilan *Spectrogram* terlihat perbedaan tingkat energi masing-masing audio. Sehingga bisa mengidentifikasi pola suara dan gangguan atau anomali dalam sinyal audio. Analisis *spectrogram* memiliki tingkat akurasi yang cukup tinggi karena apabila rekaman suara telah diubah atau dimodifikasi maka tidak berpengaruh banyak pada sebaran energi pada *spectrogram* [18]. Pada source code *Python* yang dipakai menggunakan fitur *Librosa* untuk menganalisis karakteristik *spectrogram* audio asli dengan hasil steganografi. *Librosa* adalah paket *Python* untuk pemrosesan musik dan audio yang memungkinkan pengguna untuk memuat audio di notebook sebagai array *numpy* untuk analisis dan manipulasi [19].

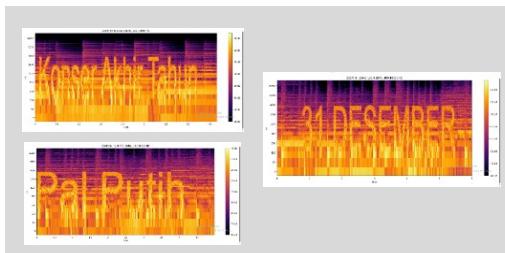


Gambar 6. *Spectrogram* hasil stego

Untuk menemukan pesan yang disisipkan di area *spectrogram*. Dilakukan *Zoom In* untuk memperbesar area *spectrogram* audio supaya dapat dengan jelas

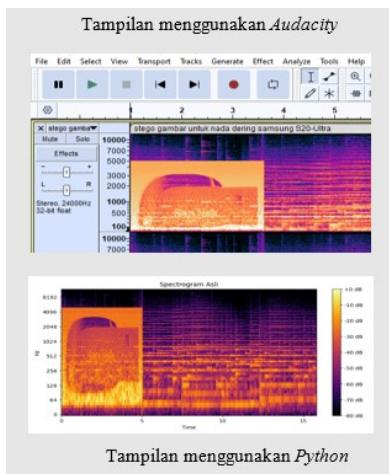


melihat isi pesan. Proses menampilkan pesan rahasia juga dilakukan menggunakan pemrograman *Python* dengan memakai fitur *Librosa* seperti pada gambar 7 di bawah ini.



Gambar 7. Zoom In Spectrogram hasil stego menggunakan Python

Implementasi steganografi dilakukan pada nada dering *Samsung Galaxy S3*. Adapun pesan yang disisipkan bertuliskan “*nama saya permadi kusuma konsentrasi forensika digital universitas islam indonesia*” dengan jumlah 10 kata, untuk tampilan pesan yang sudah disisipkan dapat dilihat menggunakan fitur *zoom in* pada aplikasi *Audacity* dan melalui *Google Colab* dengan menggunakan script Pemrograman *Python*. Selain kalimat. Peneliti juga menyisipkan pesan gambar *Flashdisk* ke area *Spectrogram* nada dering bawaan *Samsung Galaxy S20* yang bisa dilihat pada gambar 8. proses menampilkan pesan rahasia juga dilakukan menggunakan pemrograman *Python* dengan memakai fitur *Librosa*.

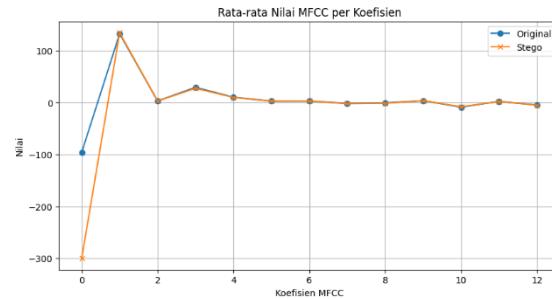


Gambar 8. Zoom In Spectrogram menampilkan gambar flashdisk

3.2. Analisis MFCC (Mel-Frequency Cepstral Coefficients)

MFCC (Mel-Frequency Cepstral Coefficients) adalah fitur audio yang mewakili karakteristik suara manusia atau ciri khas suara dalam bentuk angka-angka. *MFCC* sering digunakan dalam Pengenalan suara (*voice recognition*), analisis sinyal audio dan deteksi perbedaan atau modifikasi dalam audio. Dengan melakukan Analisis *MFCC* maka bisa menangkap perubahan kecil dalam struktur sinyal audio stego dan audio asli. Proses Analisis *MFCC* menggunakan Pemrograman *Python* di *Google Colab*.

Berikut merupakan hasil analisa *MFCC* dari audio asli lagu *Bernadya* dan hasil steganografi yang menampilkan grafik dan numerik dari kedua file.



Gambar 9. Perbandingan grafik MFCC audio asli dan steganografi lagu Bernadya

Pada grafik *MFCC* audio asli dan steganografi dari lagu *Bernadya* menunjukkan perbandingan nilai *MFCC Mean* (rata-rata koefisien *MFCC*) antara audio original dan audio stego dari lagu *Bernadya*. Dimana garis biru mewakili file audio asli, sedangkan garis oranye mewakili audio stego. Pada Sumbu X merupakan Koefisien *MFCC* dari 0 sampai 12 yang mewakili *MFCC_1* hingga *MFCC_13*. Jadi setiap titik mewakili satu koefisien. Sedangkan sumbu Y merupakan nilai rata-rata (*mean*) yang menunjukkan nilai numerik rata-rata dari masing-masing koefisien. Semakin tinggi nilainya, semakin besar karakteristik energi pada frekuensi tersebut.

Pada koefisien ke-0 (*MFCC_1*) telah terjadi perbedaan sangat besar antara Original dan Stego (turun drastis) yang mengindikasikan perubahan akibat penysipan pesan. Koefisien ke-1 (*MFCC_2*) menampilkan nilai sangat tinggi di kedua file dan sangat mirip sehingga tidak banyak berubah atau ada sedikit perubahan, sedangkan Koefisien ke-2 sampai ke-12 (*MFCC_3 - MFCC_13*) menunjukkan pola yang hampir identik hanya terjadi perubahan kecil.

Tabel 1 Nilai *MFCC Mean* lagu Bernadya

MFCC Mean	Original Mean	Stego Mean	Selisih	Keterangan
MFCC 1	-96,02	-300,23	204,20	Berubah
MFCC 2	132,44	134,43	1,99	Berubah
MFCC 3	3,00	2,91	0,09	Berubah
MFCC 4	29,28	27,36	1,92	Berubah
MFCC 5	10,27	9,85	0,42	Berubah
MFCC 6	2,72	2,61	0,11	Berubah
MFCC 7	2,79	2,89	0,10	Berubah
MFCC 8	-1,68	-2,29	0,61	Berubah
MFCC 9	-0,66	-1,44	0,79	Berubah
MFCC 10	3,42	3,60	0,18	Berubah
MFCC 11	-8,87	-8,19	0,68	Berubah
MFCC 12	2,03	1,92	0,11	Berubah
MFCC 13	-4,79	-5,18	0,39	Berubah



Kriteria penilaian dengan memberikan penilaian terhadap besar perubahan (selisih) antara audio original dan stego berdasarkan nilai *MFCC* koefisien *Mean* dan *STD*.

- Jika tidak ada perbedaan (selisih = 0), maka dianggap "Tidak berubah".
- Jika terdapat perbedaan nilai *Mean* (kurang atau lebih dari 2), maka dianggap "Berubah".

Berdasarkan gambar 8 yang menampilkan grafik *MFCC* audio asli dan steganografi dari lagu *Bernadya* telah ditemukan bahwa ada perbedaan nilai *Mean* yaitu rata-rata kekuatan suara pada tiap koefisien frekuensi yang terdapat pada audio stego dan audio asli. Tujuan Analisis *MFCC* yaitu mendeteksi perbedaan halus akibat penyisipan pesan. Nilai tersebut dapat disimpulkan memiliki selisih nilai yang berbeda yang nampak pada tabel 1 karena menunjukkan seberapa besar selisih energi suara di setiap koefisien frekuensi. Terjadinya perbedaan, karena penyisipan pesan telah mengubah struktur frekuensi, dengan adanya perbedaan nilai *MFCC* (*Mel-Frequency Cepstral Coefficients*) tersebut maka dapat digunakan untuk mendukung analisis audio steganografi.

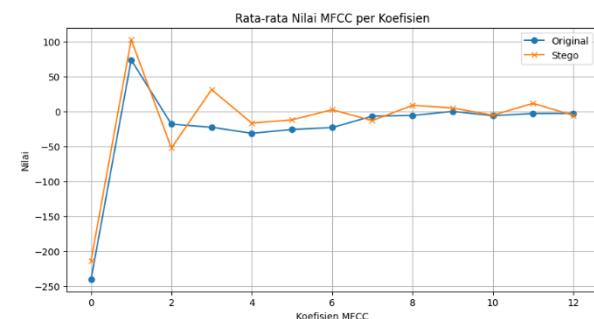
Tabel 2. Nilai MFCC STD (Standard Deviation) lagu Bernadya

Koefisien MFCC STD	Original STD	Stego STD	Selisih	Keterangan
MFCC 1	89,18	87,54	1,64	Berubah
MFCC 2	37,31	39,93	2,62	Berubah
MFCC 3	24,32	24,69	0,37	Berubah
MFCC 4	15,75	15,82	0,07	Berubah
MFCC 5	10,59	11,93	1,34	Berubah
MFCC 6	11,60	12,59	0,99	Berubah
MFCC 7	9,41	9,97	0,56	Berubah
MFCC 8	9,11	10,01	0,90	Berubah
MFCC 9	8,13	8,94	0,81	Berubah
MFCC 10	9,68	9,55	0,13	Berubah
MFCC 11	8,36	9,15	0,79	Berubah
MFCC 12	7,12	7,34	0,22	Berubah
MFCC 13	7,95	8,17	0,22	Berubah

Berdasarkan tabel 2 telah ditemukan bahwa ada perbedaan nilai *STD* (*Standard deviation*) yang menunjukkan seberapa bervariasi suara di koefisien yang terdapat pada audio stego dan audio asli. Semakin tinggi nilainya, semakin banyak perubahan atau variasi suara di area frekuensi tersebut seperti yang nampak pada tabel 2 karena menunjukkan seberapa besar selisih energi suara di setiap koefisien frekuensi. Jika audio stego dari standar deviasi lebih tinggi dari audio asli, bisa jadi ada "noise tak kasat mata" dari akibat pesan disisipkan.

Berikut merupakan Analisis perbedaan *MFCC* (*Mel-Frequency Cepstral Coefficients*) pada audio dari nada dering *Samsung Galaxy* dan setelah dilakukan proses steganografi dengan menyisipkan kalimat "*nama saya permadi kusuma konsentrasi forensika digital universitas islam indonesia*". Proses Analisis *MFCC*

menggunakan *Pemrograman Python di Google Colab* seperti pada gambar 10 dibawah ini.



Gambar 10. Perbandingan grafik MFCC audio asli dan steganografi nada dering Samsung Galaxy S3

Gambar 10 menunjukkan perbedaan nilai *MFCC* secara menyeluruh pada audio asli dan hasil stego. Pada grafik *MFCC* audio asli dan steganografi dari nada dering *Samsung* menunjukkan perbandingan nilai *MFCC Mean* (*rata-rata koefisien MFCC*) antara audio original dan audio stego dari nada dering *Samsung*. Dimana garis biru mewakili file audio asli, sedangkan garis oranye mewakili audio stego. Pada Sumbu X merupakan Koefisien *MFCC* dari 0 sampai 12 yang mewakili *MFCC_1* hingga *MFCC_13*. Jadi setiap titik mewakili satu koefisien. Sedangkan sumbu Y merupakan nilai rata-rata (*mean*) yang menunjukkan nilai numerik rata-rata dari masing-masing koefisien. Semakin tinggi nilainya, semakin besar karakteristik energi pada frekuensi tersebut.

Pada koefisien ke-0 (*MFCC_1*) telah terjadi perbedaan nilai antara Original dan Stego yang mengindikasikan perubahan akibat penyisipan pesan dengan selisih 26,91. Koefisien ke-1 (*MFCC_2*) menampilkan selisih 28,96 di kedua file sehingga mengindikasikan adanya perubahan nilai. Sedangkan Koefisien ke-2 sampai ke-12 (*MFCC_3* - *MFCC_13*) menunjukkan beberapa naik-turun antara Stego dan Original. Untuk mendapatkan analisis yang lebih mendalam, dilakukan analisis *Mean* dan *STD* pada audio asli dan hasil stego yang bisa dilihat pada tabel 4.3 dan 4.4.

Tabel 3 Nilai MFCC Mean nada dering *Samsung Galaxy S3*

Koefisien MFCC Mean	Original Mean	Stego Mean	Selisih	Keterangan
MFCC 1	-240,79	-213,87	26,91	Berubah
MFCC 2	74,13	103,09	28,96	Berubah
MFCC 3	-17,96	-52,24	34,28	Berubah
MFCC 4	-22,46	31,57	54,02	Berubah
MFCC 5	-31,06	-16,47	14,59	Berubah
MFCC 6	-25,64	-11,82	13,81	Berubah
MFCC 7	-22,92	2,70	25,63	Berubah
MFCC 8	-6,49	-12,91	6,43	Berubah
MFCC 9	-5,53	8,99	14,52	Berubah
MFCC 10	0,29	5,23	4,95	Berubah
MFCC 11	-5,85	-5,13	0,73	Berubah
MFCC 12	-2,71	11,86	14,56	Berubah
MFCC 13	-2,51	-5,38	2,87	Berubah



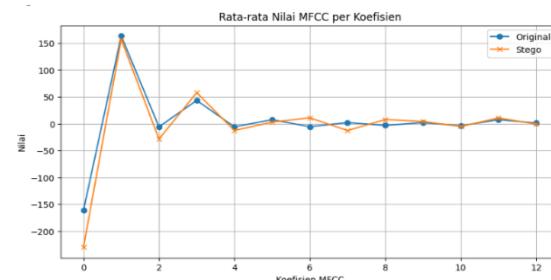
Berdasarkan gambar 8 yang menampilkan grafik MFCC audio asli dan steganografi dari nada dering *Samsung Galaxy* telah ditemukan bahwa ada perbedaan nilai Mean yaitu rata-rata kekuatan suara pada tiap koefisien frekuensi yang terdapat pada audio stego dan audio asli. Nilai tersebut dapat disimpulkan memiliki selisih nilai yang berbeda yang nampak pada tabel 3 karena menunjukkan seberapa besar selisih energi suara di setiap koefisien frekuensi. Terjadinya perbedaan nilai, karena penyisipan pesan telah mengubah struktur frekuensi, dengan adanya perbedaan nilai MFCC (*Mel-Frequency Cepstral Coefficients*) tersebut maka dapat digunakan untuk mendukung analisis audio steganografi.

Tabel 4. Nilai MFCC STD (*Standard Deviation*) nada dering *Samsung Galaxy S3*

Koefisien MFCC	Original STD	Stego STD	Selisih	Keterangan
MFCC 1	141,19	145,68	4,49	Berubah
MFCC 2	37,02	42,44	5,42	Berubah
MFCC 3	26,60	32,40	5,80	Berubah
MFCC 4	18,85	47,39	28,54	Berubah
MFCC 5	15,47	19,96	4,50	Berubah
MFCC 6	13,59	17,54	3,95	Berubah
MFCC 7	13,28	16,83	3,55	Berubah
MFCC 8	12,36	14,33	1,97	Berubah
MFCC 9	14,24	17,46	3,22	Berubah
MFCC 10	12,83	12,70	0,13	Berubah
MFCC 11	10,87	8,18	2,70	Berubah
MFCC 12	14,95	10,92	4,03	Berubah
MFCC 13	15,06	12,72	2,34	Berubah

Berdasarkan tabel 4 telah ditemukan bahwa ada perbedaan nilai *STD* (*Standard deviation*) yang menunjukkan seberapa bervariasi suara di koefisien yang terdapat pada audio stego dan audio asli. Semakin tinggi nilainya, semakin banyak perubahan atau variasi suara di area frekuensi tersebut, karena menunjukkan seberapa besar selisih energi suara di setiap koefisien frekuensi. Jika audio stego dari standar deviasi lebih tinggi dari audio asli, bisa jadi ada "noise tak kasat mata" dari akibat pesan disisipkan. Selain menyisipkan kalimat. Peneliti juga menyisipkan gambar *Flashdisk* merk *Sandisk* pada area Spectrogram nada dering *Samsung Galaxy S20*. Untuk perbandingan grafik MFCC audio asli dengan steganografi tidak jauh berbeda, ini bisa dilihat pada gambar 10 yang menunjukkan perbedaan nilai MFCC secara menyeluruh pada audio asli dan hasil stego gambar *Flashdisk* merk *Sandisk*.

Pada grafik MFCC audio asli dan steganografi dari nada dering *Samsung Galaxy S20* menunjukkan perbandingan nilai MFCC Mean (*rata-rata koefisien MFCC*) antara audio original dan audio stego dari nada dering *Samsung Galaxy S20*. Dimana garis biru mewakili file audio asli, sedangkan garis oranye mewakili audio stego. Pada Sumbu X merupakan Koefisien MFCC dari 0 sampai 12 yang mewakili MFCC_1 hingga MFCC_13.



Gambar 11. Perbandingan grafik MFCC audio asli dan steganografi gambar *Flashdisk*

Tabel 5. Nilai MFCC Mean steganografi gambar pada nada dering *Samsung S20 Galaxy S20*

Koefisien MFCC Mean	Original Mean	Stego Mean	Selisih	Keterangan
MFCC 1	-160,78	-229,66	68,89	Berubah
MFCC 2	163,08	156,60	6,48	Berubah
MFCC 3	-5,98	-28,62	22,64	Berubah
MFCC 4	43,22	57,60	14,39	Berubah
MFCC 5	-5,94	-12,75	6,81	Berubah
MFCC 6	7,42	2,88	4,54	Berubah
MFCC 7	-5,75	10,85	16,61	Berubah
MFCC 8	2,03	-12,91	14,94	Berubah
MFCC 9	-3,27	7,73	11,01	Berubah
MFCC 10	1,99	4,06	2,07	Berubah
MFCC 11	-4,04	-5,33	1,29	Berubah
MFCC 12	7,42	10,73	3,31	Berubah
MFCC 13	1,23	-0,91	2,14	Berubah

Tabel 6. Nilai MFCC STD steganografi gambar pada nada dering *Samsung S20 Galaxy S20*

Koefisien MFCC STD	Original STD	Stego STD	Selisih	Keterangan
MFCC 1	40,53	97,24	56,71	Berubah
MFCC 2	22,12	16,51	5,61	Berubah
MFCC 3	8,39	44,76	36,37	Berubah
MFCC 4	10,37	31,45	21,08	Berubah
MFCC 5	7,33	21,54	14,21	Berubah
MFCC 6	6,13	12,35	6,22	Berubah
MFCC 7	6,68	22,01	15,32	Berubah
MFCC 8	6,33	20,68	14,34	Berubah
MFCC 9	6,71	12,45	5,75	Berubah
MFCC 10	6,78	9,30	2,52	Berubah
MFCC 11	7,45	9,87	2,41	Berubah
MFCC 12	6,54	9,93	3,39	Berubah
MFCC 13	6,54	9,01	2,47	Berubah

Berdasarkan tabel 5 dan tabel 6 telah ditemukan bahwa ada perbedaan nilai Mean dan STD (*Standard deviation*) yang menunjukkan seberapa bervariasi suara di koefisien yang terdapat pada audio stego dan audio asli. Semakin tinggi nilainya, semakin banyak perubahan atau variasi suara di area frekuensi tersebut, karena menunjukkan seberapa besar selisih energi suara di setiap koefisien frekuensi. Jika audio stego dari standar deviasi lebih tinggi dari audio asli, bisa jadi ada "noise tak kasat mata" dari akibat pesan disisipkan.



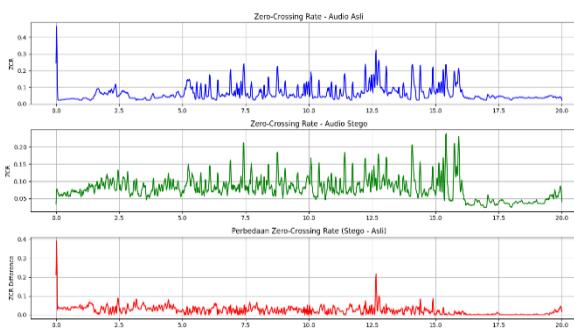
Pada analisis *MFCC* (*Mel-Frequency Cepstral Coefficients*) dari audio steganografi lagu *Bernadya* yang terdapat pesan "konser akhir tahun", "pal putih" dan "31 desember", nada dering *Samsung Galaxy* yang terdapat pesan "nama saya permadi kusuma konsentrasi forensika digital universitas islam indonesia" dan nada dering *Samsung Galaxy S20* yang disisipkan gambar *Flashdisk* merk *Sandisk* dengan audio asli secara menyeluruh. Ditemukan nilai *MFCC Mean* dan *STD* dengan hasil yang berbeda dengan audio aslinya. Sehingga pada analisis *MFCC* ini disimpulkan bahwa audio asli dan hasil stego mengalami perbedaan akibat adanya proses steganografi.

3.3. Analisis ZCR (Zero-Crossing Rate)

Analisis ini didasarkan pada seberapa sering sinyal audio melintasi garis nol dalam satuan waktu. Dengan membandingkan *ZCR* antara audio asli dan audio yang telah disisipi pesan, maka dapat melihat apakah ada peningkatan atau pola yang tidak biasa. Jika karakteristik audio asli dan hasil stego tersebut menunjukkan tingkat perbedaan yang besar, maka dapat disimpulkan bahwa *Analisis Zero-Crossing Rate* dari audio asli dan hasil stego adalah berbeda. Proses Analisis *ZCR* menggunakan Pemrograman *Python* di *Google Colab*.



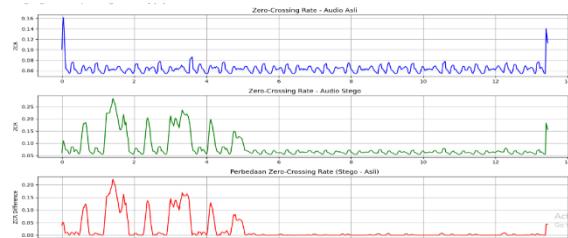
Gambar 12. Grafik ZCR audio asli dan steganografi Bernadya



Gambar 13. Grafik ZCR audio asli dan steganografi nada dering Samsung S3

Pada tahapan ini, analisis *ZCR* di gunakan untuk melihat pola umum audio asli dengan hasil stego dari lagu *Bernadya*. Grafik pertama (biru) menunjukkan *ZCR* dari audio asli, kedua (hijau) dari audio stego, dan ketiga (merah) mewakili perbandingan dari audio asli dan stego. Jika ada lonjakan atau pola perbedaan yang tidak biasa di grafik ketiga, itu bisa

mengindikasikan adanya penyisipan data. Sehingga pada tahapan ini kan jelas terlihat tingkat suatu energi dari masing-masing audio.



Gambar 14. Grafik ZCR audio asli dan steganografi nada dering Samsung Galaxy S20

Pada analisis *ZCR* (*Zero-Crossing Rate*) dari audio yang dianalisis yaitu lagu *Bernadya*, nada dering *Samsung Galaxy*, dan nada dering *Samsung Galaxy S20* dengan audio stego secara menyeluruh, ditemukan perbandingan grafik *ZCR* dengan hasil yang berbeda dengan audio aslinya. Sehingga pada analisis *ZCR* (*Zero-Crossing Rate*) disimpulkan bahwa audio asli dan hasil stego mengalami perbedaan akibat adanya proses steganografi.

3.4 Hasil pengujian penyisipan informasi rahasia ke dalam audio

Berikut merupakan tabel hasil penyisipan informasi rahasia ke dalam lagu *Bernadya* dengan *Samsung*.

Tabel 7. Hasil penyisipan informasi rahasia ke dalam audio

Isi informasi rahasia	Audio penampung (mp3)	Size audio asli	Size audio stego	Hasil
Konser Akhir Tahun, Pal Putih, 31 Desember.	Lagu Bernadya - Hidup Harus Tetap Berjalan.mp3	7.01 mb	33.6 mb	✓
Nama saya Permadi Kusuma konsentrasi forensika digital Universitas Islam Indonesia	Nada Dering Samsung Galaxy S3.mp3	317 kb	3.36 mb	✓
Nada Dering Samsung Galaxy S20 Ultra.mp3	263 kb	1.23 mb		✓

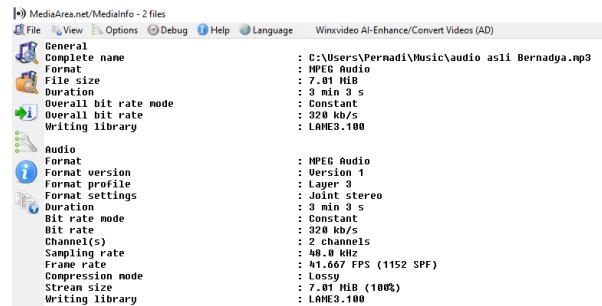
Pada Tabel 7 menampilkan ukuran audio setelah dilakukan proses stego dapat menambah size lagu. Hasil yang didapat menunjukkan penyisipan pesan rahasia telah berhasil, namun terdapat perubahan visual dari *MFCC* dan *ZCR*. Untuk mengetahui penyebab audio mengalami perubahan size maka dilakukan analisis meta data.



3.5. Analisis Metadata

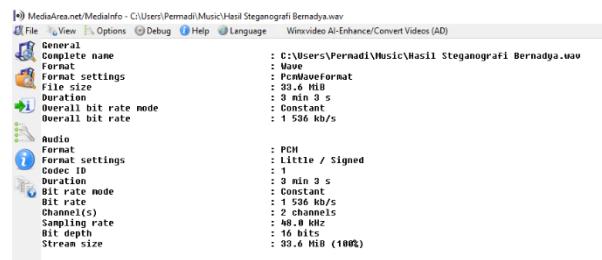
Analisis metadata audio asli dan hasil stego dilakukan guna melihat perbedaan metadata dari kedua file dikarenakan terdapat penambahan size setelah dilakukan proses steganografi. Analisis metadata menggunakan *Software Mediainfo*. Metadata yang tampil berisi informasi rinci seperti format file, sample rate, durasi, *bitrate*, dan channel, yang semuanya dapat memengaruhi ukuran akhir file. Dengan membandingkan metadata sebelum dan sesudah penyisipan pesan, peneliti dapat mengidentifikasi apakah peningkatan ukuran disebabkan oleh perubahan pada parameter teknis seperti peningkatan *bitrate*, penambahan channel, atau penggunaan format encoding yang berbeda (dari MP3 ke WAV).

3.5.1. Analisis metadata lagu Bernadya



Gambar 15. Metadata audio lagu Bernadya

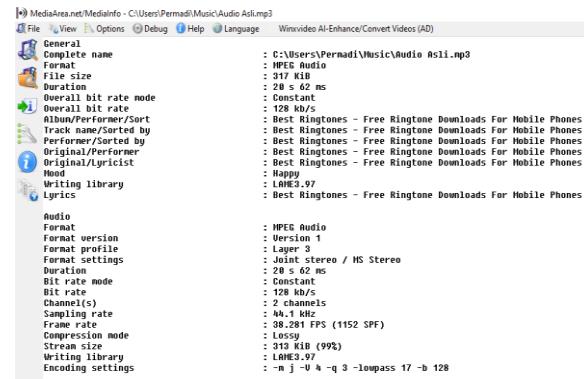
Pada gambar 15 menampilkan metadata dari file lagu *Bernadya* yang belum dilakukan proses steganografi. Size audio memiliki ukuran 7.01 mb dengan durasi 3 menit 3 detik. *Overall bit rate* 320 kb/s. Pada gambar 16 menampilkan metadata dari file lagu *Bernadya* yang sudah dilakukan proses steganografi. Size audio memiliki ukuran 33.6 mb dengan durasi 3 menit 3 detik. *Overall bit rate* 1536 kb/s.



Gambar 16. Metadata audio stego lagu Bernadya

3.5.2. Analisis metadata nada dering Samsung Galaxy S3

Pada gambar 17 menampilkan metadata dari file nada dering *Samsung Galaxy S3* yang belum dilakukan proses steganografi. Size audio memiliki ukuran 311 kb dengan durasi 20 detik. *Overall bit rate* 128 kb/s.



Gambar 17. Metadata nada dering audio asli Samsung Galaxy S3

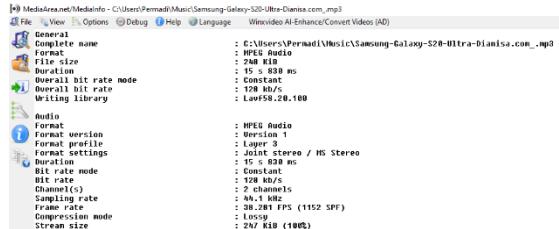
Pada gambar 18 menampilkan metadata dari file nada dering *Samsung* yang sudah dilakukan proses steganografi. Hasilnya file stego yang terdapat pesan rahasia dalam bentuk teks memiliki size audio 3.37 mb dengan durasi 20 detik. *Overall bit rate* 1411 kb/s.



Gambar 18. Metadata nada dering Samsung Galaxy S3 hasil stego.

3.5.3. Analisis metadata nada dering Samsung Galaxy S20

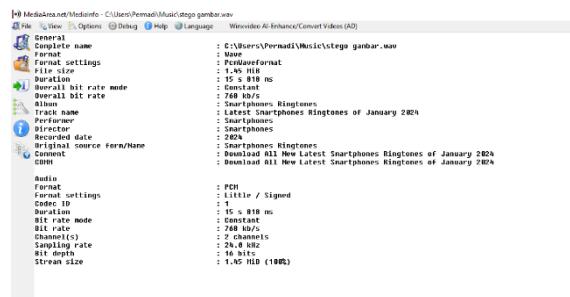
Pada gambar 19 menampilkan metadata dari file nada dering *Samsung Galaxy S20* yang belum dilakukan proses steganografi. Size audio memiliki ukuran 248 kb dengan durasi 15 detik. *Overall bit rate* 128 kb/s.



Gambar 19. Metadata nada dering Samsung Galaxy S20 audio asli

Pada gambar 20 menampilkan metadata dari file nada dering *Samsung Galaxy S20* yang sudah dilakukan proses steganografi. Hasilnya file stego yang terdapat pesan rahasia dalam bentuk gambar memiliki size audio 1.45 mb dengan durasi 15 detik. *Overall bit rate* 768 kb/s.





Gambar 20. Metadata nada dering Samsung Galaxy S20 audio stego

Dari analisis metadata audio asli dan stego dapat diketahui bahwa proses steganografi audio melalui *Spectrogram* dapat menambah *Overall bit rate* audio Lagu *Bernadya* yang awalnya 320 kb/s menjadi 1536 kb/s, nada dering *Samsung Galaxy S3* 128 kb menjadi 1411 kb/s, dan nada dering *Samsung Galaxy S20* 248 kb/s menjadi 768 kb/s. Menurut sumber dari *Wikipedia*, standar *bitrate* untuk lagu berformat *MP3* bervariasi antara 32 kbps hingga 320 kbps, tergantung pada kualitas kompleksitas audio. *Bitrate* yang umum digunakan untuk musik adalah 128 kbps sebagai kualitas minimum, 192–256 kbps untuk kualitas sedang, dan 320 kbps sebagai bitrate tertinggi yang mendekati kualitas CD. Sedangkan pada buku [20] dengan judul “*Diminishing returns of higher mp3 bit*” menjelaskan bahwa sebagian besar pendengar audio tidak dapat membedakan antara file *MP3* dengan *bitrate* 128 kbps dan audio tanpa kompresi. Namun, pada bitrate 48 kbps, perbedaan kualitas menjadi jelas. Ini menunjukkan bahwa 128 kbps dianggap sebagai titik optimal antara kualitas dan ukuran file untuk audio umum, termasuk nada dering.

Adanya penambahan bitrate pada audio stego dikarenakan adanya perubahan format audio *MP3* ke *WAV*. Format *WAV* menyimpan data lebih banyak per detik dibanding *MP3*, sehingga *WAV* membawa lebih banyak data per detik dibandingkan *MP3*. Selain itu, analisis *MFCC* (*Mel-Frequency Cepstral Coefficients*) nilai *Mean* dan Standar Deviasi menampilkan grafik dan nilai numerik audio asli dengan hasil stego yang berbeda yaitu terdapat perubahan pada setiap koefisien. Pada analisis *ZCR* (*Zero-Crossing Rate*) juga menunjukkan perbandingan pada grafik *ZCR* audio asli dan stego, sehingga dapat disimpulkan bahwa ukuran audio bertambah setelah dilakukan proses steganografi dikarenakan adanya perubahan format audio misalnya *MP3* ke *WAV* dan juga adanya penyatuan file stego yang digabung dengan file audio asli.

3.6. Tabel Hasil Pengujian Audio Stego

Pada hasil pengujian audio stego dengan audio asli pada tabel 8 menunjukkan audio yang dilakukan proses stego akan mengubah size audio. Begitupun dengan karakteristik nilai ZCR, MFCC dan metadata audio. Pendekatan berlapis seperti spectrogram,

MFCC, ZCR, dan metadata digunakan karena masing-masing memberi informasi berbeda untuk mendeteksi perubahan pada audio. Spectrogram menunjukkan perubahan visual, MFCC melihat karakter suara, ZCR mendeteksi gangguan halus, dan metadata memeriksa perubahan teknis file. Teknik masking menyisipkan pesan di frekuensi yang tidak sensitif bagi telinga manusia, dengan memperhatikan bit-depth audio agar perubahan tidak terdengar. Toleransi error juga digunakan agar pesan tetap bisa diambil meskipun terjadi sedikit gangguan. Pendekatan ini membuat analisis lebih akurat dan menyeluruh.

Tabel 8. Hasil penyisipan informasi rahasia ke dalam audio

No	Audio	Size audio asli	Size audio stego	Hasil	Pesan
1	Bernadya - Untungnya Hidup Harus Tetap Berjalan.mp3	7.01 mb	33.6 mb	Berubah	Terbaca
2	Samsung Galaxy S20.mp3	317 kb	3.36 mb	Berubah	Terbaca
3	Samsung Galaxy 3.mp3	263 kb	1.23 mb	Berubah	Terbaca
4	Alan Walker - The Spectre	3.5 mb	6.7 mb	Berubah	Terbaca
5	Alan Walker - Fade	3.7 mb	7.4 mb	Berubah	Terbaca
6	Ariel Noah - menghapus jejakmu	4.5 mb	7.5 mb	Berubah	Terbaca
7	Ariel Noah - Sahabat	3.3 mb	5.4 mb	Berubah	Terbaca
8	Paramore - Still Into You	4.5 mb	7.6 mb	Berubah	Terbaca
9	Paramore - decode	4.7 mb	6.2 mb	Berubah	Terbaca
10	Paramore - All I Wanted	6.5 mb	5.5 mb	Berubah	Terbaca

4. Kesimpulan

Berdasarkan hasil yang didapat pada proses seganografi dengan menggunakan teknik Masking dapat ditarik kesimpulan yaitu implementasi steganografi ke dalam *Spectrogram* audio lagu *Bernadya*, nada dering *Samsung Galaxy S3* dan nada dering *Samsung Galaxy S20* menggunakan teknik Masking melalui *Spectrogram* audio telah berhasil. Teknik masking terbukti dapat diimplementasikan secara efektif untuk menyisipkan pesan tersembunyi



ke dalam spectrogram audio. Berdasarkan analisis steganografi audio yang dilakukan menunjukkan bahwa *Spectrogram* memperlihatkan pola visual tambahan yang tidak terdapat pada audio asli. Penelitian juga melakukan pengujian terhadap user. Namun user tidak dapat mendeteksi pesan rahasia tersebut tanpa menggunakan tools atau software. Dikarenakan pesan rahasia disembunyikan pada area *spectrogram* audio yang hanya bisa dilihat menggunakan software khusus seperti audacity, praat dan pemrograman python.

Parameter akustik seperti *MFCC (Mel-Frequency Cepstral Coefficients)* dan *ZCR (Zero-Crossing Rate)* juga menunjukkan perbedaan nilai rata-rata dan standar deviasi yang berbeda, sehingga mengindikasikan adanya modifikasi spektrum akibat proses penyisipan. Selain itu, metadata file audio asli dengan stego menunjukkan perubahan size dan Overall bit rate pada audio stego. Sehingga hasil analisis menunjukkan karakteristik audio asli berbeda dibandingkan dengan hasil stego. Pada teknik Masking yang diimplementasikan pada media audio masih terdapat kekurangan yaitu hasil analisis menunjukkan bahwa karakteristik audio asli dengan audio stego memiliki perbedaan pada gelombang suara, size audio dan *Overall bit rate*.

Reference

