

STEGANALISIS AUDIO DIGITAL DENGAN DISCRETE WAVELET TRANSFORM MENGGUNAKAN METODE SUPPORT VECTOR MACHINE CLASSIFIER

AUDIO DIGITAL STEGANALYSIS WITH DISCRETE WAVELET TRANSFORM USING SUPPORT VECTOR MACHINE CLASSIFIER

Nanang Adi Setyawan¹, Dr. Ir. Bambang Hidayat, DEA.², Sofia Sa'idah, S.T., M.T.³

Program Studi S1 Teknik Telekomunikasi, Fakultas Teknik, Universitas Telkom

nanangadi93@gmail.com¹, bhidayat@telkomuniversity.ac.id², sofiasaidahsfi@gmail.com³

Abstrak

Dokumen audio sebagai *cover* dalam menyembunyikan dokumen lain dengan cara disisipkan menggunakan teknik steganografi merupakan hal yang sudah cukup banyak ditemukan untuk melindungi kerahasiaan pesan seseorang. Dalam pembacaannya dan pendeteksiannya diperlukan suatu teknik untuk mengidentifikasi apakah ada suatu pesan dalam suatu dokumen audio, hal ini disebut dengan steganalisis. Steganalisis mencakup pencarian keberadaan suatu pesan dalam medium tertentu sampai dengan mengetahui isi pesan yang disisipkan dalam suatu medium lain.

Dalam penelitian ini dibuat suatu perangkat lunak yang mampu mendeteksi keberadaan suatu pesan tersembunyi berupa teks dalam sebuah file dengan format WAVE. Metode yang digunakan untuk mendeteksi yaitu dengan discrete wavelet transform dengan dikalsifikasikan dengan support vector machine classifier.

Hasil dari steganalisis ini berupa akurasi sistem sebesar 90%. Penelitian ini diharapkan dapat digunakan lagi bagi pihak-pihak yang berkepentingan dan mempunyai kebutuhan mengenai informasi dan pesan rahasia seperti kepolisian.

Kata Kunci : *Steganalisis, Steganografi, Discrete Wavelet Transform, Support Vector Machine, WAVE.*

Abstract

The audio document as a cover in hiding other documents by inserting using steganography techniques is something that is already quite commonly found to protect the confidentiality of a person's message. In its readings and its detection required a technique to identify whether there is a message in an audio document, this is called steganalisis. Steganalisis involves searching for the existence of a message in a particular medium to know the contents of a message inserted in another medium.

In this research, a software that can detect the existence of a hidden message of text in a file with WAVE format. The method used to detect that is with discrete wavelet transform with dikalsifikasikan with support vector machine classifier.

The results of this steganalisis can determine the accuracation system 90%. This study is expected to be used again for interested parties and has a need for confidential information and messages such as the police.

Keywords: *Steganalisis, Steganography, Discrete Wavelet Transform, Support Vector Machine, WAVE.*

1. Pendahuluan

1.1 Latar Belakang

Perkembangan ilmu pengetahuan dan teknologi yang sangat pesat berpengaruh terhadap berbagai aspek dan bidang salah satunya adalah tentang teknologi pertukaran informasi. Dalam bertukar informasi pada zaman sekarang ini sudah banyak diberi kemudahan, pada mulanya pertukaran informasi hanya menggunakan suatu media konvensional dan hanya berupa teks, untuk saat ini sudah semakin beragam dalam segi media maupun informasi yang bisa disampaikan. Saat ini informasi yang bisa disampaikan tidak hanya berupa teks, tetapi bisa berupa teks, video, gambar, maupun audio yang pengirimannya juga tidak memakan waktu yang lama.

Kemudahan dalam pertukaran informasi ini harus diimbangi dengan teknik untuk mengamankan informasi yang disampaikan, agar tidak diterima ataupun terbaca oleh orang yang salah, salah satu metodenya adalah metode steganografi. Steganografi merupakan suatu teknik pengamanan pesan dengan menyembunyikan pesan ke dalam suatu media tertentu atau dalam *cover-object*. Akan tetapi dari sisi penerima, orang tersebut harus bisa memahami apakah dalam sebuah *cover-object* terdapat suatu pesan rahasia atau tidak. Teknik untuk bisa mengetahui apakah terdapat suatu pesan atau tidak dalam *cover* disebut dengan steganalisis. Steganalisis memberikan solusi bagi koresponden untuk saling berkomunikasi dengan baik, agar baik dari sisi pengirim yang telah mengirim file pesan rahasia yang disisipkan dalam sebuah *cover* bisa diketahui oleh penerima. Steganalisis bisa dilakukan di berbagai media *cover*, antara lain citra, audio, maupun video.

Penelitian terkait steganalisis yang telah dilakukan sebelumnya adalah steganalisis dengan menggunakan metode *support vector machine radial basis function (SVM Rbf) classifier* (Tajrij 2009). Penelitian tersebut menghasilkan keakuratan yang cukup tinggi mencapai 63,3 % untuk mendeteksi informasi tersembunyi pada suatu berkas audio.

Pada tugas akhir ini akan dibahas tentang steganalisis dengan *discrete wavelet transform* untuk mendeteksi apakah ada suatu pesan tersembunyi, file yang akan digunakan yaitu file audio dengan format *.wav*. Dalam pengerjaan tugas akhir ini akan dilakukan steganalisis dengan *discrete wavelet transform* menggunakan metode *support vector machine classifier* untuk klasifikasinya.

1.2 Rumusan Masalah

Pada Tugas Akhir ini terdapat beberapa rumusan masalah yang akan dibahas, diantaranya:

1. Bagaimana merancang simulasi sistem steganalisis audio digital dengan *discrete wavelet transform* menggunakan metode *support vector machine classifier*.
2. Bagaimana mengidentifikasi ada atau tidaknya pesan tersembunyi dalam suatu berkas audio dengan format *.wav*.
3. Bagaimana tingkat akurasi steganalisis menggunakan metode *support vector machine classifier*.

1.3 Tujuan Penelitian

Tujuan dari pembuatan tugas akhir ini adalah :

1. Merancang suatu simulasi sistem steganalisis audio digital dengan *discrete wavelet transform* dengan menggunakan metode *support vector machine classifier* pada file audio dengan format *.wav*.
2. Mengidentifikasi ada atau tidaknya pesan tersembunyi pada suatu berkas audio dengan format **.wav*.
3. Mengetahui tingkat akurasi steganalisis menggunakan metode *support vector machine classifier*.

1.4 Batasan Masalah

1. Metode steganalisis yang digunakan adalah dengan *discrete wavelet transform* dengan metode *support vector machine classifier*.
2. Sistem yang dirancang difokuskan untuk mengidentifikasi apakah dalam suatu file audio terdapat pesan tersembunyi atau tidak, tidak sampai dengan ekstraksi pesan.
3. Tugas akhir ini difokuskan pada membahas steganalisis dengan data audio digital dengan format *.wav*.
4. Tidak membahas dan menganalisis mengenai proses dan metode steganografi yang digunakan pada file stego yang dideteksi.
5. Tugas akhir ini menggunakan aplikasi Matlab R2009a.

1.5 Metode Penelitian

Metode penelitian yang digunakan dalam pengerjaan tugas akhir ini adalah :

1. Studi literatur, dengan mempelajari buku-buku referensi yang berkaitan dengan obyek terkait dan melalui situs internet yang mendukung dalam penulisan tugas akhir ini.
2. Tahap Eksperimen, yaitu tahapan dimana dilakukan percobaan dalam membuat aplikasi simulasi untuk mengidentifikasi pesan rahasia dari file audio dengan menerapkan konsep-konsep yang telah diperoleh dari tahapan sebelumnya.
3. Tahap Uji dan Analisis, yaitu dilakukan pengujian terhadap hasil yang didapatkan sehingga sesuai dengan harapan serta melakukan analisis dari aplikasi yang dibuat.
4. Diskusi dengan pembimbing mengenai hasil analisis data-data yang diperoleh dari hasil studi literatur. Selain itu berguna untuk mendiskusikan pemecahan masalah atas kesulitan yang dihadapi.
5. Melakukan perancangan sistem.
6. Melakukan pengujian sistem dan analisis terhadap hasil uji penelitian yang dilakukan.

7. Pengambilan kesimpulan terhadap hasil analisis dan pembuatan laporan Tugas akhir dari seluruh kegiatan penelitian.

2. Dasar Teori

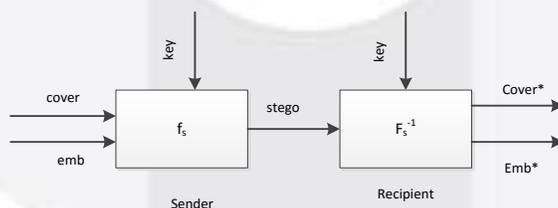
2.1 Steganografi^[5]

Steganografi adalah seni dan pengetahuan dalam menyembunyikan pesan. Suatu sistem steganografi mampu menyembunyikan pesan ke dalam suatu *cover* media sehingga tidak akan menimbulkan suatu kecurigaan jika terdapat pesan rahasia di dalamnya. Di masa lalu orang-orang biasa menggunakan tinta tak tampak (*invisible ink*) untuk berkomunikasi pada steganografi, tetapi dengan teknologi sekarang bisa memberi kemudahan untuk menyembunyikan pesan di suatu media dengan berbagai metode dan *cover*.

Steganografi berasal dari Bahasa Yunani, yaitu "*stegos*" yang artinya tersembunyi dan "*graphien*" yang artinya tulisan, jadi steganografi secara harfiah mempunyai arti tulisan tersembunyi. Dalam steganografi media yang digunakan untuk *cover* bisa bermacam-macam, umumnya media yang digunakan dapat berupa teks, citra, audio, dan video.

Steganografi yang merupakan teknik untuk menyembunyikan pesan mempunyai beberapa perbedaan jika dibandingkan dengan kriptografi, perbedaannya dapat dilihat dari hasil keluarannya. Pada kriptografi biasanya data keluaran yang dihasilkan biasanya tampak seolah-olah berantakan, tetapi itu bisa dikembalikan seperti pesan semula, sedangkan pada steganografi keluaran yang dihasilkan akan berbeda, yaitu akan sama dengan bentuk asli media atau *cover*. Sehingga dengan menggunakan steganografi ini tidak mudah diketahui jika terdapat pesan di dalamnya.

Proses steganografi secara umum menggunakan kunci sebagai sarana untuk membuka pesan yang disisipkan di dalam media. Dalam steganografi terdiri atas dua proses yaitu proses *embedding* dan proses ekstraksi. Proses *embedding* adalah proses menyisipkan pesan ke dalam media *cover*. Sedangkan proses ekstraksi yaitu proses memperoleh kembali pesan yang telah disisipkan dari media *cover*. Dalam steganografi bisa juga ditambahkan enkripsi dan dekripsi untuk menambah keamanan data. Untuk system steganografi pada umumnya dapat dilihat pada gambar 2.1 sebagai berikut :



Gambar 2.1 Sistem Steganografi

Dalam steganografi mempunyai beberapa istilah yang merupakan properti dari steganografi, antara lain :

- *Embedded-message* : pesan yang disembunyikan
- *Cover-object* : media yang digunakan untuk menyembunyikan pesan
- *Stego-object* : media yang sudah berisi *embedded-message*
- *Stego-key* : kunci yang digunakan untuk menyimpan pesan pada *cover-object* dan mengekstraksi pesan dari *stego-object*
- *Steganalysis* : proses untuk mendeteksi keberadaan pesan rahasia dalam sebuah *file*

Dalam steganografi, memanfaatkan keterbatasan indra manusia, dalam hal ini keterbatasan dalam pendengaran dan pengelihatn manusia, karena dalam steganografi hasil keluaran sistem mempunyai persepsi yang sama dengan masukan aslinya. Berdasarkan pengertian diatas maka *audio stegography* memiliki arti suatu teknik menyembunyikan pesan atau informasi tertentu ke dalam suatu file audio tanpa diketahui keberadaannya.

2.2 Steganalisis^[5]

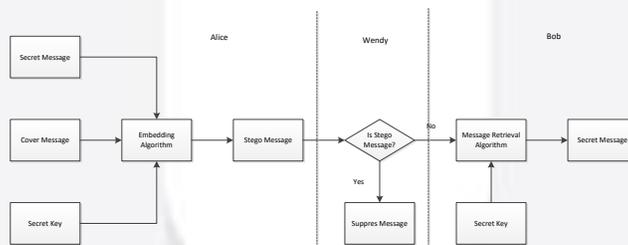
Steganalisis merupakan teknik yang digunakan untuk melawan keberadaan steganografi dengan memungkinkan adanya metode deteksi, ekstraksi, destruksi, dan manipulasi dari data tersembunyi yang ada di dalam *stego object*. Steganalisis ini bisa dilakukan dengan berbagai macam cara, salah satunya hanya ditujukan untuk

mendeteksi keberadaan suatu data yang tersembunyi. Dalam keyataannya, teknik atau metode yang digunakan dalam steganalisis secara empiris berkisar antara lain:

1. Menganalisa dari perubahan yang dilakukan terhadap data file.
2. Menganalisa dari ciri-ciri file yang telah menggunakan software tertentu untuk steganografi.
3. Membandingkan file asli, lalu dicari perbedaannya dan pola apa yang digunakan sehingga cara ini bukan saja dapat diketahui file telah mengalami proses steganografi dapat pula diketahui pesan yang disembunyikan.

Penyembunyian informasi dengan menggunakan media elektronik membutuhkan perubahan pada sifat dari media itu sendiri. Hal ini menyebabkan beberapa bentuk degradasi atau karakteristik yang tidak biasa dari media yang disisipi informasi. Karakteristik seperti ini dapat menjadi tanda bahwa terdapat penyisipan pesan sehingga tujuan dari steganografi pun menjadi patah.

Serangan dan analisis terhadap informasi yang tersembunyi dapat dilakukan dalam beberapa bentuk, seperti menemukan, menyadap, dan melumpuhkan atau menghancurkan informasi tersembunyi tersebut. Bentuk-bentuk ini juga bervariasi tergantung pada metode yang digunakan untuk menyisipkan informasi ke dalam media penutup. Tujuannya adalah bukan untuk menghilangkan atau melumpuhkan informasi yang tersembunyi seperti hak cipta, tetapi untuk menunjukkan pendekatan-pendekatan yang mudah diserang dan memanfaatkannya untuk menginvestigasi informasi tersembunyi. Proses steganalisis secara umum dapat dilihat pada gambar 2.2 sebagai berikut :



Gambar 2.2 Proses steganalisis secara umum

2.3 Audio Digital Format .WAV^[6]

Format file WAVE adalah bagian dari RIFF spesifikasi Microsoft digunakan untuk penyimpanan file multimedia. Sebuah file RIFF dimulai dengan file header diikuti dengan urutan potongan data (*data chunk*). Sebuah file WAVE sering hanya file RIFF dengan *single* "WAVE" *chunk* yang terdiri dari dua *sub-chunk*. berikut gambar dari struktur file dengan format .wav.

endian	File offset (bytes)	field name	Field Size (bytes)	Description
big	0	ChunkID	4	The "RIFF" chunk descriptor
little	4	ChunkSize	4	
big	8	Format	4	
big	12	Subchunk1ID	4	
little	16	Subchunk1Size	4	The "fmt" sub-chunk describes the format of the sound information in the data sub-chunk
little	20	AudioFormat	2	
little	22	NumChannels	2	
little	24	SampleRate	4	
little	28	ByteRate	4	
little	32	BlockAlign	2	
little	34	BitsPerSample	2	
big	36	Subchunk2ID	4	The "data" sub-chunk Indicates the size of the sound information and contains the raw sound data
little	40	Subchunk2Size	4	
little	44	data	Subchunk2Size	

Gambar 2.3 Bagian Struktur file .WAV ^[6]

Keterangan bagian struktur WAV adalah sebagai berikut:

1. Bagian *Main Chunk*
 - ChunkID: berisi kata "RIFF" dalam format ASCII.

- ChunkSize: berisi informasi ukuran chunk.
 - Format: berisi kata "WAVE".
2. Bagian *Format Chunk*
 - SubChunk1ID: berisi kata "fmt".
 - SubChunk1Size: berisi informasi ukuran subchunk1
 - AudioFormat: informasi jenis kompresi data chunk. Misalnya bernilai 1 untuk kompresi PCM.
 - NumChannels: banyaknya channel. Misal: Mono=1, Stereo=2.
 - SampleRate: sample rate dari file wav, misal 8000 untuk 8000Hz, 44100 untuk 44100 Hz.
 - ByteRate: banyaknya byte tiap detik. $ByteRate = SampleRate \times NumChannels \times BitsPerSample / 8$.
 - BitsPerSample: ukuran bits untuk tiap sampel. Misal: 8 bit = 8
 3. Bagian *Data Chunk*
 - SubChunk2ID: berisi kata "data".
 - SubChunk2Size: berisi informasi ukuran subchunk2. $SubChunk2Size = NumSamples \times NumChannels \times BitsPerSample / 8$. $NumSamples = (DataByte / NumChannels) / BitsPerSample$.
 - Data: Data suara aktual dalam byte, merepresentasi-kan amplitudo tiap sampel dari sinyal.

2.4 Discrete Wavelet Transform

Wavelet merupakan bentuk pengklasifikasian sinyal dengan melakukan penskalaan dan pergeseran masing-masing dari sinyal sumber yang mempunyai noise. Discrete wavelet transform merupakan salah satu bentuk diskrit dari sinyal wavelet yang bisa digunakan untuk menghilangkan noise. Pada proses DWT akan mengklasifikasikan sinyal sumber menjadi dua klasifikasi yaitu frekuensi tinggi dengan resolusi waktu yang tinggi dan frekuensi rendah dengan resolusi frekuensi yang tinggi.

Proses *discrete wavelet transform* mempunyai beberapa keunggulan jika dibandingkan dengan transformasi fourier , yaitu mampu menentukan amplitudo saat frekuensi sampling yang diberikan tidak tepat, DWT juga dinilai lebih sensitive dalam menunjukkan perubahan bentuk sinyal pada rentang frekuensi tertentu dan mampu mendeteksi tren dan kemiripan dari beberapa sinyal yang sedang diproses.

Prinsip dasar dari *discrete wavelet transform* yaitu menentukan bagaimana cara mendapatkan representasi waktu dan skala dari sebuah sinyal dengan teknik pemfilteran digital dan operasi sub-sampling^[7]. Sinyal informasi dilewatkan pada *high-pass filter* dan *low-pass filter* , kemudian dilakukan proses dekomposisi satu tingkat yaitu dengan mengambil setengah dari masing-masing keluaran yang nantinya akan dianggap sebagai sample melalui operasi sub-sampling. Proses dekomposisi dapat dituliskan dengan persamaan sebagai berikut :

$$Y_{high} [k] = \sum_{in} x[n]h[2k - n] \tag{2.1}$$

$$Y_{low} [k] = \sum_{in} x[n]g[2k - n] \tag{2.2}$$

Keterangan :

$Y_{high} [k]$ = hasil dari *high-pass filter*

$Y_{low} [k]$ = hasil dari *low-pass filter*

$x[n]$ =sinyal asal

$h[n]$ =*high-pass filter*

$g[n]$ =*low-pass filter*

Setelah melalui tahapan dekomposisi, koefisien *wavelet* dibentuk kembali melalui proses rekonstruksi yang disebut dengan *inverse discrete wavelet transform (IDWT)*, sehingga didapatkan persamaan sebagai berikut :

$$x[n] = \sum_k (y_{high}[k]h[-n + 2k] - y_{low}[k]g[-n + 2k]) \tag{2.3}$$

Proses rekonstruksi dimulai dengan menggabungkan koefisien DWT pada akhir dekomposisi dengan sebelumnya melakukan upsampling melalui *highpass filter* dan *lowpass filter*.

Secara umum, penyisipan pesan dilakukan dengan cara memodifikasi koefisien pada rentang frekuensi LL, LH, HL, dan HH yang merupakan rentang frekuensi hasil dekomposisi menggunakan *wavelet* dimana data pesan dianggap sebagai rangkaian bilangan w dengan panjang L yang disisipkan pada koefisien rentang yang dipilih f . Sehingga pada proses steganalisis harus melakukan klasifikasi data stego ke dalam frekuensi tinggi dan frekuensi rendah sebagai koefisien wavelet^[9].

2.5 Analisis Statistik

Pengertian dari analisis statistik adalah metode untuk memberi gambaran atas suatu kejadian melalui bentuk-bentuk yang sederhana dapat berupa angka maupun grafik sehingga kunci keberhasilan dari analisis statistik terletak pada pemakaiannya. Dalam penelitian ini digunakan analisis deskriptif yaitu bagian dari ilmu statistika yang mengolah serta menyajikan data tanpa mengambil kesimpulan untuk populasi, atau dengan kata lain hanya melihat gambaran secara umum dari data yang didapatkan^[8].

Data statistic diperoleh dari hasil survey acak serta tidak terorganisir dengan baik sehingga data yang terkumpul harus diringkas dalam bentuk table maupun grafik yang dapat digunakan sebagai acuan untuk proses pengambilan keputusan. Penyajian tabel dalam analisis deskriptif dapat berupa presentasi grafis seperti histogram, pie chart, dan lain lain. Dalam menyajikan data statistic juga diperlukan ukuran eksak yang disebut dengan *summary statistics* yang digunakan dalam pengambilan keputusan yaitu :

1. *Central tendency*, seperti *mean*.
2. Ukuran disperse, seperti standar deviasi.

2.5.1 Mean

Statistik berupa rata-rata(mean) merupakan nilai yang dianggap cukup representative untuk menggambarkan nilai yang terdapat dalam data yang bersangkutan. *Mean* sebagai nilai sentral yang digunakan sebagai pengukuran lokasi suatu distribusi frekuensi.

Bila data sample terdiri dari sejumlah nilai-nilai hasil pengamatan yang tidak terlalu besar, maka akan langsung dicari dari data yang bersangkutan tanpa harus menyusun ke dalam distribusi frekuensi, sehingga rata-rata hitung adalah penjumlahan dari nilai-nilai dibagi dengan jumlah total observasi sebesar n . persamaan untuk menghitung mean adalah sebagai berikut :

$$E(x) = \frac{1}{n} \sum_{k=1}^n x_k \quad (2.4)$$

Dengan n =jumlah data ; x =nilai data; $k=1,2,3,\dots,n$

2.5.2 Standar Deviasi

Pengukuran varians menurut Karl Pearson dirumuskan sebagai berikut :

$$Var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - E(x))^2 \quad (2.5)$$

Standarisasi unit pengukuran diatas dilakukan melalui proses pengakaran yaitu :

$$s = \left(\frac{1}{n} \sum_{i=1}^n (x_i - E(x))^2 \right)^{1/2} \quad (2.6)$$

2.6 Support Vector Machine Classifier^[1]

Dalam pengerjaan tugas akhir ini akan diteliti *filtering* sinyal audio tanpa adanya *watermark* yang menyebabkan perubahan pada metrik kualitas yang berbeda dengan *embedding* sinyal audio. Untuk mengklasifikasikan suatu sinyal audio akan digunakan metode *support vector machine polynomial*. Klasifikasi ini digunakan untuk mengklasifikasikan apakah sinyal yang dideteksi merupakan sinyal stego atau bukan.

Metode *support vector machine* merupakan metode yang didasarkan pada optimalisasi fungsi multidimensional. Metode *support vector machine* ini digunakan untuk mengklasifikasikan suatu unit ke dalam grup yang cocok dengan fitur-fitur yang dimiliki oleh unit tersebut. Metode SVM ini merupakan perkembangan dari metode *linear regression classifier*, yang mempunyai prinsip dasar mirip seperti *linear classifier*. *Support vector machine* akan menemukan *hyperplane* yang terbaik pada input *space* kemudian selanjutnya dikembangkan agar dapat bekerja pada masalah non linear dengan memasukkan konsep *kernel trick* pada ruang kerja dengan dimensi tinggi sehingga metode ini bisa mengoptimalkan fungsi pada beberapa dimensi untuk melakukan pengambilan keputusan.

Dalam *support vector machine* akan meminimalisasi resiko empiris yang masih berupa set data latih. Untuk fitur data latih $(f_i, g_i), i = 1, \dots, N$, $g_i \in [-1,1]$, fitur vektor F pada *hyperplane* diberikan dengan $w^T + b = 0$, dimana w adalah bidang normal. Set vektor akan terpisah secara optimal jika set vektor ini terpisah tanpa eror dan jarak antara vektor terdekat dengan *hyperplane* pada posisi jarak maksimal.

$$g_i [(wF_i) + b] \geq 1, i = 1, 2, \dots, N \quad (2.7)$$

Sehingga jarak dari $d(w,b;F)$ dari *hyperplane* (w,b) ke vektor F adalah

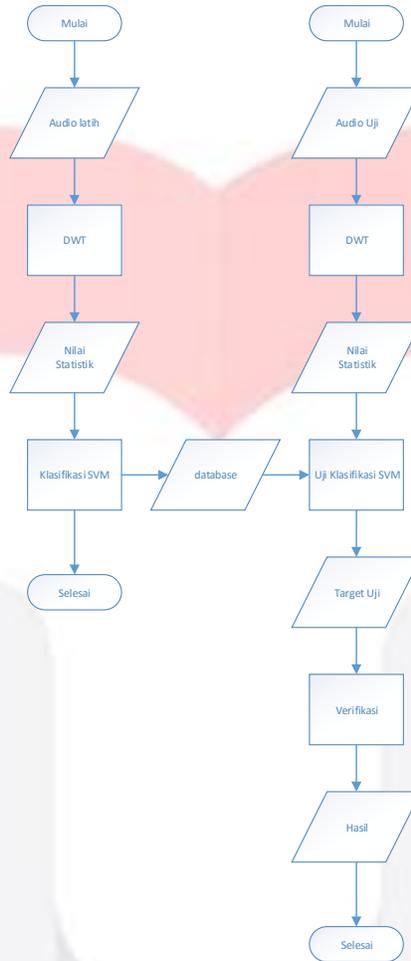
$$d(w,b;F) = \frac{|w^T F + b|}{\|w\|} \quad (2.8)$$

hyperplane optimal akan diperoleh dengan memaksimalkan margin.

3. Pembahasan

3.1 Desain Perancangan Sistem Steganalisis

Pada penelitian ini desain perancangan sistem steganalisis yang digunakan ditunjukkan pada gambar 3.2. sebagai berikut :



Gambar 3.1 Konfigurasi Blok Perancangan Audio Latih dan Audio Uji

3.2 Lingkup Pengujian

Pengujian pada tugas akhir ini menggunakan metode support vector machine untuk mendeteksi keberadaan pesan tersembunyi pada media audio. Media audio yang dideteksi adalah berkas audio dengan format *.wav*. Frekuensi dari file audio adalah 10KHz dengan panjang 16 bit. Durasi file audio adalah selama lima detik dengan informasi pesan yang disisipkan adalah berupa pesan teks. Penyisipan dilakukan dengan menggunakan Matlab dengan metode LSB.

3.3 Skenario Pengujian Sistem

Pengujian dilakukan untuk mengetahui performansi sistem yang telah dibuat. Performansi sistem dapat dilihat berdasarkan banyaknya audio stego yang terdeteksi dengan benar dengan jumlah audio stego yang diujikan. Banyaknya data audio digital yang diujikan yaitu terdiri dari 10 data audio asli yang belum disisipi pesan dengan 20 data audio stego yang telah disisipi pesan masing-masing terdiri dari 10 data audio disisipi pesan sebanyak 18 karakter dan 10 data audio disisipi pesan sebanyak 28 karakter. Skenario yang digunakan dalam tugas akhir ini yaitu melakukan ekstraksi ciri dengan metode DWT untuk memperoleh data statistic berupa *mean*, *skewness*, dan standar deviasi kemudian melakukan klasifikasi dengan SVM *quadratic*, kemudian menghitung nilai akurasi sistem dari audio uji dengan beberapa parameter pengujian yaitu :

1. Pengujian analisis pengaruh banyaknya huruf terhadap ciri statistic data.
2. Pengujian analisis pengaruh banyaknya huruf terhadap tingkat akurasi.
3. Pengujian analisis pengaruh level DWT terhadap tingkat akurasi.
4. Pengujian analisis pemilihan fitur ciri untuk klasifikasi.

3.4 Hasil Pengujian

Berdasarkan skenario pengujian yang telah ditetapkan sebelumnya, maka dilakukan analisis sebagai berikut

3.4.1 Analisis Pengaruh Banyaknya Huruf Terhadap Ciri Statistik Data

Dalam penelitian ini dilakukan analisis banyaknya huruf yang disisipkan terhadap ciri statistik data (mean dan standar deviasi) dari data uji. Audio asli disisipkan pesan menurut jumlah karakter dengan menggunakan teknik penyisipan yaitu LSB pada audio dengan durasi 5 detik dan 10 detik. Pada tabel 4.1 dipaparkan isi pesan yang tersembunyi dalam audio stego.

Tabel 4.1 Pesan yang disisipkan pada audio stego

Jumlah huruf	Isi pesan
18 Karakter	i love you so much
28 Karakter	Because love is never enough

Setelah dilakukan penyisipan, dilakukan pengujian pada audio asli dan audio stego untuk mencari ciri statistik pada masing-masing audio yaitu berupa mean dan standar deviasi. Ciri statistik audio stego dan audio asli berdasarkan jumlah huruf yang disisipkan dapat dilihat dari tabel 4.2 berikut :

Tabel 4.2 Ciri statistik audio digital

Metode Steganografi	Jenis audio	Ciri	
		Mean	Standar deviasi
LSB	Asli	2.1777342e-05	0.351609
	18 karakter	1.5398e-05	0.253498
	28 karakter	1.5397e-05	0.253398

Dari tabel yang diperoleh dapat dilihat perbedaan ciri antara audio asli dan audio stego. Pada audio stego LSB, semakin banyak karakter yang disisipkan nilai standar deviasi semakin kecil. Terlebih jika dibandingkan dengan audio asli tanpa disisipi pesan, nilainya cukup terlihat berbeda.

3.4.2 Analisis Pengaruh Banyak Huruf Terhadap Tingkat Akurasi

Dalam penelitian ini dilakukan analisis jumlah karakter yang disisipkan terhadap akurasi sistem. Audio asli disisipkan pesan menurut jumlah karakter seperti yang dipaparkan pada tabel 4.1. penyisipan pesan teks dilakukan dengan teknik LSB pada audio dengan durasi 5 detik dan 10 detik. Audio kemudian dilakukan proses DWT dan diklasifikasikan menurut kelasnya berdasarkan ciri statistiknya dengan menggunakan metode SVM. Keberhasilan dan kebenaran dari klasifikasi sistem ditunjukkan dengan nilai akurasi yang diperoleh. Tingkat akurasi berdasarkan jumlah huruf pada file stego dengan menggunakan uji DWT pada audio stego yang disisipi dengan metode LSB ditunjukkan pada tabel 4.3.

Tabel 4.3 Pengaruh banyaknya huruf (karakter) terhadap akurasi

Metode Steganografi	Jumlah Karakter	Akurasi
LSB	18 karakter	90%
	28 karakter	90%

Dari tabel diatas, jumlah karakter yang disisipkan tidak terlalu mempengaruhi tingkat akurasi yang diperoleh . Hal ini disebabkan perbedaan jumlah karakter yang disisipkan tidak terlalu signifikan sehingga tidak terlalu mempengaruhi hasil dan akurasinya. Disebabkan juga karena jumlah sample yang diambil sangat sedikit sehingga kurang merepresentasikan kondisi sebenarnya. Pada akurasi diperoleh cukup tinggi karena pada metode steganografi LSB ciri data audio asli tidak banyak diubah sehingga tidak banyak terdapat perbedaan pada akurasi yang diperoleh.

4.3.3 Analisis pengaruh level DWT terhadap tingkat akurasi

Dalam penelitian ini dilakukan analisis pengaruh level DWT terhadap tingkat akurasi sistem. Setelah audio disisipkan pesan berupa teks dengan menggunakan metode LSB, dilakukan metode DWT beberapa level . Pada saat pengujian , dapat dilihat beberapa perbedaan ciri terhadap penggunaan level DWT, perubahan ciri yang diperoleh mempengaruhi besarnya ukuran akurasi dalam melakukan klasifikasi kebenaran kelas audio uji tersebut.

Akurasi sistem dari pengujian masing-masing level untuk audio uji menggunakan DWT ditunjukkan pada gambar 4.3 sebagai berikut.

Tabel 4.4 Akurasi sistem steganalisis pada masing-masing level DWT

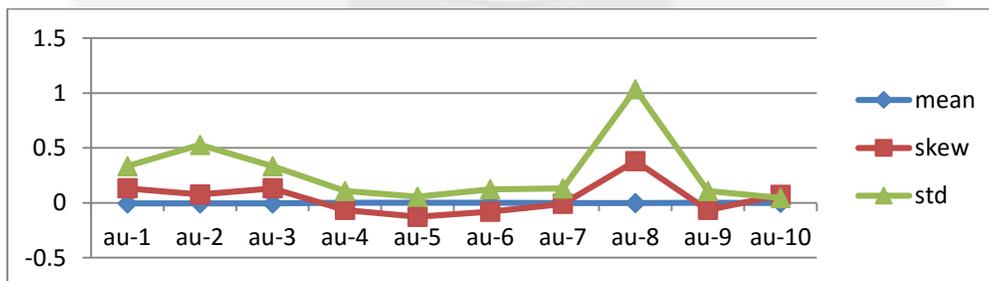
Metode Steganografi	Level	18 karakter	28 karakter
LSB	1	30%	30%
	2	30%	30%
	3	20%	20%
	4	70%	70%
	5	90%	90%
	6	80%	80%

Dari tabel diatas, dapat dilihat pengaruh banyak level terhadap akurasi sistem, yaitu pada penggunaan level DWT semakin tinggi maka semakin besar tingkat akurasi yang diperoleh, tetapi hal ini terbatas hanya sampai level 5. Pada level 6 kembali diperoleh penurunan akurasi sistem. Pada level 1 sampai level 5, semakin tinggi level DWT maka semakin tinggi nilai akurasi yang diperoleh, hal ini terjadi karena penggunaan level mempengaruhi banyak perbedaan ciri statistik yang didapat, sehingga semakin banyak perbedaan ciri yang didapat, semakin bisa dibedakan antara audio asli dengan audio stego. Tetapi pada penelitian ini terbatas hanya sampai level 5. Saat dilakukan percobaan dengan level DWT pada level 6 ditemukan penurunan akurasi, hal ini dikarenakan semakin panjang proses dekomposisi DWT, semakin banyak juga sinyal informasi yang hilang diakibatkan dari proses *downsampling* sehingga sinyal yang diproses sudah tidak lagi akurat.

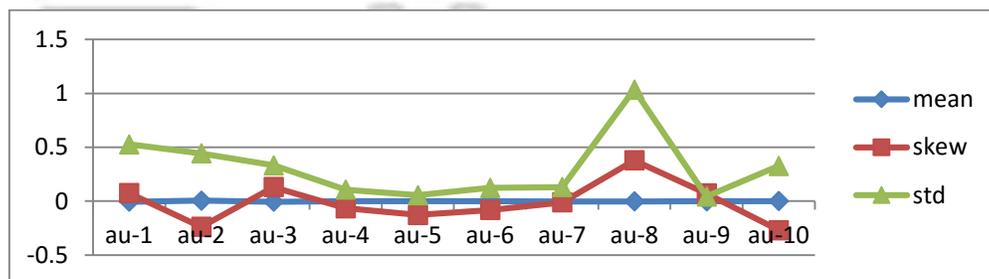
4.3.4 Analisis pemilihan fitur ciri untuk klasifikasi

Pada penelitian ini dilakukan analisis pemilihan ciri data statistik yang digunakan untuk klasifikasi. Ciri yang digunakan untuk klasifikasi pada penelitian ini adalah mean, skewness, dan standar deviasi dari masing-masing data uji. Audio asli yang telah disisipkan pesan berupa teks dengan menggunakan teknik steganografi menggunakan metode LSB pada audio durasi 5 detik dan 10 detik.

Grafik dari masing-masing data statistik yang digunakan pada pengujian ini ditunjukkan pada gambar 4.4. Data statistik berikut merupakan data statistik yang diperoleh dari pengujian dengan menggunakan DWT pada level 5.



(a)



(b)

Gambar 4.4 Data statistik *mean*, *skewness*, dan standar deviasi pada audio uji dengan DWT level 5

Dari grafik diatas diperoleh kesimpulan bahwa untuk melakukan pengujian menggunakan data statistik untuk melakukan klasifikasi dapat digunakan berbagai macam nilai statistik. Untuk nilai statistik yang paling baik untuk digunakan sebagai ciri adalah mean karena untuk masing-masing audio uji yang berbeda, nilainya mempunyai simpangan yang paling kecil, sehingga jika digunakan untuk klasifikasi akan diperoleh nilai yang akurat.

4. Kesimpulan

Berdasarkan hasil perancangan dan analisa sistem steganalisis yang telah digunakan, maka dapat diambil kesimpulan sebagai berikut :

1. Steganalisis pada file audio dengan format *.wav dapat dilakukan dengan membandingkan nilai karakteristiknya dilihat dari nilai mean dan varian nya.
2. Support Vector Machine dapat digunakan untuk melakukan klasifikasi pada proses steganalisis pada penelitian ini. Yaitu untuk melakukan klasifikasi apakah suatu pesan tersebut telah tersisipi pesan lain atau tidak.
3. Pada pengujian dilakukan skenario pengaruh jumlah karakter terhadap akurasi. Hasil yang diperoleh yaitu jumlah karakter yang disisipkan tidak terlalu mempengaruhi tingkat akurasi yang diperoleh hanya terdapat kenaikan sangat sedikit. Hal ini disebabkan perbedaan jumlah karakter yang disisipkan tidak terlalu signifikan sehingga tidak terlalu mempengaruhi hasil dan akurasinya. Pada akurasi diperoleh kenaikan juga karena pada metode steganografi LSB ciri data audio asli diubah sehingga akan terdapat perbedaan pada akurasi yang diperoleh.
4. Pada pengujian DWT diperoleh ciri berupa mean dan standar deviasi. Hasil ciri yang diperoleh adalah terdapat perbedaan ciri statistic yang cukup jauh pada audio asli dan audio stego. Sedangkan pada audio stego dengan perbedaan jumlah karakter yang disisipkan juga terdapat perbedaan namun tidak signifikan.
5. Dari keseluruhan pengujian yang dilakukan dengan metode DWT, performansi sistem yang diperoleh adalah sebesar 95 %. Pengujian dilakukan dengan menggunakan metode DWT dan klasifikasi menggunakan SVM kernel quadratic.

Daftar Pustaka :

- [1] A. S. Nugroho, A. B. Witarto, and D. Handoko, "Support Vector Machine," 2003.
- [2] C. Taswell and C. Toolsmiths, "The What , How , and Why of Wavelet Shrinkage Denoising A Simple Explanation and a 1-D Example," pp. 1–11, 1998.
- [3] H. Ozer, "Steganalysis of Audio Based on Audio Quality Metrics Steganalysis of Audio Based on Audio Quality Metrics ," no. August 2015, 2003.
- [4] K. Tajrij, "*Steganalisis Pada Media Audio Menggunakan Metode Support Vector Machine Radial Basis Function (SVM-RBF) Classifier*", Skripsi, Jurusan Ilmu Komputer FMIPA, Bogor,2009.
- [5] R. Bohme, "Principles of Modern Steganography and Steganalysis," pp. 11–78, 2010.
- [6] T. Wave, "WAVE PCM soundfile format."
- [7] Elizabeth Juwita, "Simulasi dan Steganalisis Audio Digital Menggunakan Metode Discrete Wavelet Transform dan Principal Component Analysis",2014.
- [8] M.Saini and R.Chhikara,"DWT Feature based Blind Image Steganalysis using Neural Network Classifier,"vol 4,no.4,pp. 776-782, 2015.
- [9] Visoline Ivaprida, "Steganalisis Audio Format WAV Menggunakan Metode Discrete Wavelet Transform dan Linear Discriminant Analysis", 2016.
- [10] Vivin Fauziah, "Steganalisis Untuk File Audio Berformat MP3 Dengan Metode Least Significant Bit Pada Klasifikasi Principal Component Analysis (PCA)". 2017.
- [11] Sheira Banu, "Steganalisis Citra Digital Berbasis Discrete Cosine Transform Menggunakan K-Nearest Neighbor". 2017.
- [12] H.Hafid, A. Islamiyati, P.S. Statistia, U. Hasanuddin, I. Kepercayaan, and G. Bumi, " Interval Kepercayaan

Skewness dan Kurtosis Menggunakan Bootstrap pada Data Kekuatan Gempa Bumi". 2015.

- [13] Visoline Ivaprilda S, "Steganalisis Audio Format WAV Menggunakan Metode Discrete Wavelet Transform dan Linear Discriminant Analysis". 2016.



Telkom
University