

Teknik *Steganography* menggunakan Algoritma Genetika

Miranti Andhita Scantya^{1,a)}, Triati Dewi Kencana Wungu^{2,b)} dan Suprijadi^{3,c)}

¹Program Studi Magister Sains Komputasi,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Bandung, Jl. Ganesha no. 10 Bandung,
Indonesia, 40132

²Kelompok Keilmuan Fisika Nuklir dan Biofisika, Program Studi Fisika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Bandung, Jl. Ganesha no. 10 Bandung,
Indonesia, 40132

³Kelompok Keilmuan Fisika Teori Energi Tinggi dan Instrumentasi, Program Studi Fisika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Bandung, Jl. Ganesha no. 10 Bandung,
Indonesia, 40132

^{a)}mirantya22@gmail.com (corresponding author)

^{b)}triati@fi.itb.ac.id

^{c)}supri@fi.itb.ac.id

Abstrak

Steganography adalah pendekatan proteksi data yang prosesnya adalah dengan menyembunyikan pesan atau informasi dalam suatu media tanpa merusak kualitas media tersebut. Dengan teknik *steganography* ini proses transaksi informasi diharapkan akan menjadi lebih aman dan terlindungi dari pihak-pihak yang tidak berhak mengaksesnya. Media yang digunakan dalam *steganography* bisa berupa citra digital, teks, audio, maupun video tanpa menunjukkan perubahan yang nyata dalam kualitas media yang digunakan. Pada paper ini pesan atau informasi yang disisipkan pada citra adalah file berbentuk teks. Untuk menyisipkan pesan digunakan teknik algoritma genetika. Teknik ini digunakan untuk membandingkan kualitas citra saat disisipi pesan dengan algoritma genetika dan tanpa teknik apapun. Kualitas citra setelah proses *steganography* selesai akan diukur menggunakan PSNR (Peak Signal to Noise Ratio) dan MSE (Mean Square Error). Hasil akhir pada pengujian ini menunjukkan bahwa pesan berhasil disembunyikan dan teknik *steganography* menggunakan algoritma genetika, dilihat dari hasil PSNR, lebih baik dibandingkan teknik *steganography* dengan metode yang digunakan pada paper pembandingan. Hasil PSNR terbesar dari pengujian adalah 66,79.

Kata Kunci: *Steganography*, Algoritma Genetika, Peak signal to noise ratio

PENDAHULUAN

Saat ini untuk melakukan pertukaran informasi semakin mudah dan cepat dengan adanya perkembangan teknologi. Informasi yang dikirim dan juga media untuk pengiriman informasi bisa berupa teks, citra, audio maupun video. Namun pada perkembangan teknologi ini selain mendapatkan dampak positif, terdapat juga dampak negatif dari pertukaran informasi tersebut. Contohnya yang berkaitan dengan keamanan data atau pesan yang dikirim. Jika tidak ada metode untuk mengamankan pesan yang dikirim maka informasi didalam pesan tersebut bisa dicuri dan diakui kepemilikannya oleh orang lain. Maka dari itu diciptakan metode untuk menyembunyikan data, salah satunya adalah *steganography*. *Steganography* adalah seni dan ilmu untuk menyembunyikan data yang dapat berupa teks, audio, citra, dan video dalam sebuah citra yang belum memiliki pesan [1].

Percobaan steganography dengan berbagai teknik sudah banyak dilakukan misalnya dalam paper karya Elham Ghasemi, Jamshid Shanbehzadeh, dan Nima Fassih berjudul "High Capacity Image Steganography Using Wavelet Transform and Genetic Algorithm". Dalam Paper tersebut dilakukan percobaan steganography dengan teknik GA menggunakan transformasi wavelet (GADWT) untuk membandingkan hasil percobaan berupa nilai PSNR dengan hasil PSNR dari paper sebelumnya.

Dalam paper ini, teknik steganography yang digunakan adalah genetic algorithm dan discrete cosine transforms berdasarkan (GADCT). Penerapan GA di steganography dapat meningkatkan optimasi penyisipan pesan dan kualitas citra. Kualitas citra setelah disisipkan pesan diukur menggunakan Peak Signal to Noise Ratio (PSNR).

METODE KOMPUTASI

Steganography

Steganography adalah seni dan ilmu untuk menyembunyikan informasi. Pada sistem steganography konten yang tersembunyi disisipkan dalam cover media sehingga tidak menimbulkan kecurigaan pada orang-orang yang tidak berkaitan. Di masa lalu, orang menggunakan tato tersembunyi atau tinta tak terlihat untuk menyampaikan konten steganography. Saat ini teknologi komputer dan jaringan menyediakan saluran komunikasi yang mudah digunakan untuk steganography. Pada dasarnya, proses menyembunyikan informasi dalam sistem steganography dimulai dengan mengidentifikasi bit redundant cover media (yang dapat dimodifikasi tanpa merusak medium integritasnya) [1].

Ada tiga aspek berbeda pada sistem menyembunyikan informasi: capacity, security dan robustness [1]. Capacity atau kapasitas mengacu pada banyak informasi yang dapat disembunyikan dalam media. Security atau keamanan mengacu pada seberapa besar kemampuan seseorang untuk mendeteksi pesan tersembunyi. Sedangkan robustness mengacu pada seberapa banyak modifikasi yang dapat diterima cover media.

Algoritma Genetika

Algoritma genetika atau Genetic algorithm (GA) pertama kali dipublikasikan oleh John Holland pada tahun 1975 di Amerika Serikat. Pada saat itu, GA memiliki bentuk yang sangat sederhana sehingga disebut Simple GA (SGA). Ciri dari SGA adalah SGA lebih menitik beratkan pada rekombinasi (cross-over). Pada saat inialisasi GA membangkitkan sejumlah individu sebagai anggota populasi, secara acak atau berdasarkan suatu pengetahuan tertentu. Jumlah individu dalam populasi tersebut selalu tetap selama proses evolusi. Setiap individu hanya memiliki satu kromosom dengan panjang yang tetap selama proses evolusi. Artinya jumlah gen dalam setiap kromosom tidak akan bertambah atau berkurang selama proses evolusi. Setiap gen hanya bisa bernilai 1 atau 0 yang pada saat inialisasi dibangkitkan secara acak dengan probabilitas yang sama. Setiap kromosom didekodekan dan kemudian dievaluasi sehingga diperoleh nilai fitness. Nilai fitness digunakan sebagai parameter dalam pemilihan orangtua yang bersifat proporsional terhadap nilai fitness. Dua kromosom orangtua yang terpilih akan direkombinasi menggunakan metode N-point atau uniform untuk menghasilkan dua kromosom baru (anak). Pada kedua kromosom anak tersebut mungkin saja terjadi mutasi yang berupa pembalikan bit. Setelah mempunyai kromosom baru yang mempunyai jumlah yang sama dengan kromosom lama, maka kromosom lama digantikan dengan kromosom baru. Dengan bergantinya semua kromosom, maka populasi tersebut dikatakan telah menjalani proses evolusi untuk satu generasi [4].

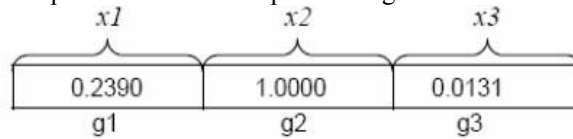
Proses Evolusi

Berikut adalah proses evolusi pada SGA[4]:

- Skema pengkodean

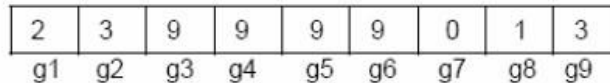
Masing-masing kromosom berisi sejumlah gen yang mengkodekan informasi yang disimpan didalam kromosom.. Individu-individu dikodekan kedalam kromosom- kromosom. Terdapat tiga skema yang umum digunakan dalam pengkodean [4]:

- Real-number encoding. Pada skema ini, nilai gen berada dalam interval $[0,R]$. Dimana R adalah biulangan real positif. Contoh terdapat dalam gambar dibawah ini.



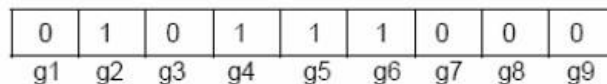
Gambar 1. Real-number encoding

- Discrete decimal encoding. Setiap gen bisa bernilai salah satu bilangan bulat dalam interval $[0,9]$. Contoh terdapat dalam gambar dibawah ini.



Gambar 2. Discrete decimal encoding

- Binary encoding. Setiap gen hanya bisa bernilai 0 atau 1. Contoh terdapat dalam gambar dibawah ini.



Gambar 3. Binary encoding

1. Evaluasi fungsi fitness

Fungsi fitness adalah suatu persamaan yang disesuaikan dengan permasalahan untuk mencapai solusi yang benar dan optimal. Fungsi fitness digunakan untuk mengevaluasi suatu individu. Persamaan fungsi fitness secara umum adalah:

$$f = h \tag{1}$$

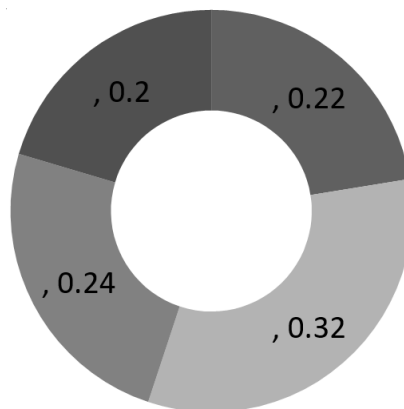
$$f = \frac{1}{(h+a)} \tag{2}$$

Pada persamaan pertama, fungsi fitness adalah fungsi objektif. Persamaan ini digunakan jika nilai fungsi objektif menghasilkan nilai maksimal. Sebaliknya jika fungsi objektif menghasilkan nilai minimum maka persamaan dipakai adalah persamaan kedua, suatu individu dievaluasi dengan melihat nilai fitness terbaik atau nilai fitness yang paling maksimal dimana α adalah bilangan yang dianggap sangat kecil untuk pembagian dengan nilai 0. Agar setiap individu tidak memiliki nilai fitness yang sama dilakukan linear fitness ranking yang bertujuan untuk melakukan penskalaan nilai-nilai fitness [4]. Berikut persamaan linear fitness ranking:

$$f(i) = - \left(f_{\max} - (f_{\max} - f_{\min}) \left[\frac{R(i)-1}{N-1} \right] \right) \tag{3}$$

2. Seleksi Orang tua

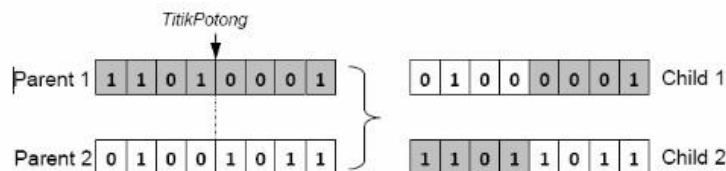
Pemilihan dua kromosom sebagai orang tua dilakukan secara proporsional berdasarkan nilai-nilai fitness-nya. Nilai fitness dihitung dengan fungsi fitness yang bergantung pada masalah yang sedang dikerjakan. Salah satu algoritma seleksi yang umum digunakan adalah roulette-wheel. Algoritma ini menirukan permainan roulette-wheel dimana kromosom-kromosom menempati potongan lingkaran pada roulette secara proporsional sesuai dengan nilai fitness-nya. Kromosom dengan fitness terbesar menempati ruang yang besar di roulette-wheel sehingga kemungkinan kromosom itu terpilih juga besar. Namun komponen terlemah masih mempunyai kesempatan untuk dipilih walaupun dengan kemungkinan yang kecil. Contoh Roulette Wheel terdapat dalam gambar dibawah ini.



Gambar 4. Contoh Roulette wheel

3. Rekombinasi

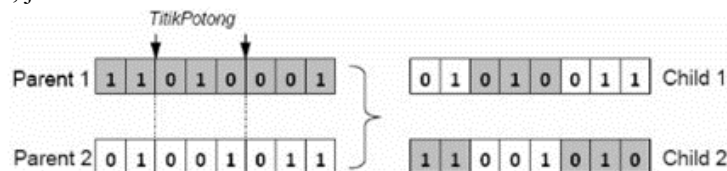
a) Setelah dua kromosom terpilih sebagai orang tua, maka kedua kromosom tersebut akan direkombinasi untuk menghasilkan dua anak. Terdapat tiga jenis rekombinasi yang biasa digunakan pada SGA, yaitu: Rekombinasi satu titik (1-point cross-over) Pada rekombinasi satu titik, suatu titik rekombinasi dibangkitkan, misalkan titik berada pada gen ke-2. Maka setiap gen diluar titik tersebut disilang, sehingga menghasilkan dua anak yang mempunyai gen yang bersilangan.



Gambar 5. Metode silang dengan one point cross over

b) Rekombinasi banyak titik (N-point cross-over)

Pada rekombinasi ini, jumlah titik rekombinasi bisa lebih dari 2.



Gambar 6. Metode silang dengan 2 point cross over.

c) Rekombinasi seragam (Uniform cross-over)

Pada rekombinasi ini, rekombinasi dilakukan berdasarkan suatu pola acak dimana penamaan gen diganti oleh 1 jika pewarisan gen menyilang dan 0 jika pewarisan gen tidak menyilang. Atau pun sebaliknya.

4. Mutasi

Mutasi pada SGA mungkin terjadi secara acak dan bebas pada setiap gen. Proses terjadinya mutasi dilakukan dengan membangkitkan bilangan acak di interval [1,0]. Jika bilangan acak lebih kecil dari probabilitas mutasi (Pm), maka gen tersebut dimutasi dengan cara membalik nilai biner gen tersebut. Probabilitas mutasi ditentukan dengan 1/NL sampai 1/L. Dimana N adalah jumlah kromosom dalam populasi dan L adalah jumlah gen dalam satu kromosom. [6]. Contoh mutasi kromosom dapat dilihat pada gambar 7,8 dan 9.



Gambar 7. Mutasi pada kromosom



Gambar 8. Mutasi pada Gen



Gambar 9. Mutasi level bit

5. Seleksi Survivor

Seleksi survivor dilakukan dengan mengganti kromosom lama dengan kromosom baru.

Peak Signal to Noise Ratio (PSNR)

Perhitungan PSNR digunakan untuk membandingkan kualitas citra hasil dengan citra asal. Semakin tinggi nilai PSNR maka semakin bagus kualitas citra tersebut.

Di bawah ini adalah persamaan untuk rumus PSNR:

$$PSNR = 10 \times \left(\log_{10} \left(\frac{255^2}{MSE} \right) \right) \tag{4}$$

Dengan MSE:

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (f_{ij} - g_{ij})^2 \tag{5}$$

Dimana M dan N adalah ukuran panjang dan lebar citra, f_{ij} menunjukkan pixel cover image dan g_{ij} adalah pixel stego image.

HASIL DAN PEMBAHASAN

Perancangan Sistem

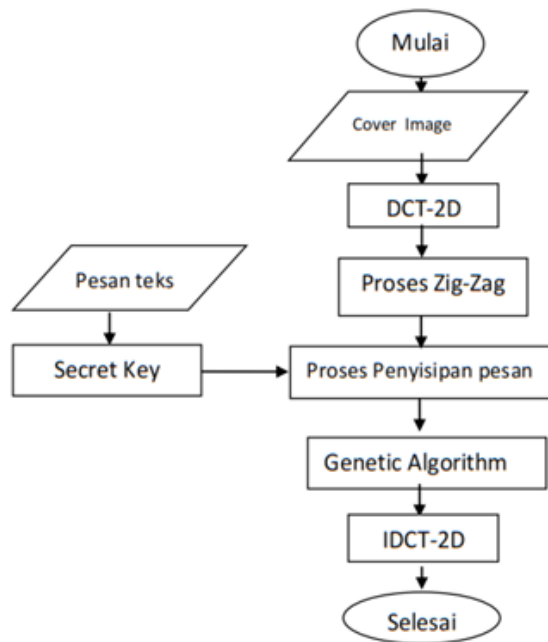
1.1 Perancangan Sistem

Sistem yang dibangun memiliki 2 proses. Yaitu proses penyisipan pesan pada citra dan ekstrasi pesan pada citra untuk mendapatkan bit-bit pesan yang telah disisipkan kedalam citra. Berikut adalah diagram alir yang menjelaskan tentang proses penyisipan pesan atau Embedding process serta proses ekstrasi pesan atau extraction algorithm.

Embedding Algorithm :

Bagi Cover image kedalam Blok 4x4. Citra dengan ukuran 512x512 di bagi kedalam blok- blok berukuran 4x4 untuk memudahkan proses DCT.

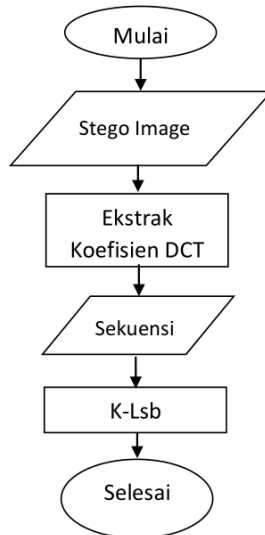
2. Cari representasi domain frekuensi dari blok-blok tersebut dengan DCT dan dapatkan 4 subbands yang terdiri dari LL1, HL1, LH1, dan HH1 melalui proses zig-zag.
3. Generate kromosom dengan 16 gen yang berisi nomor pixel dari setiap blok 4x4 sebagai fungsi mapping.
4. Masukkan bit pesan di koefisien DCT tiap pixel berdasarkan fungsi mapping. Untuk memilih nilai dari k, citra di evaluasi dari k=3 sampai 6. k yang bernilai 1 atau 2, menghasilkan kapasitas menyembunyikan yang rendah dengan kualitas visual tinggi dari stego image dan k yang bernilai 7 atau 8, menghasilkan kapasitas menyembunyikan yang tinggi dan kualitas visual yang rendah.
5. Algoritma Genetika diterapkan pada saat penyisipan pesan. Evaluasi fitness dilakukan untuk memilih fungsi mapping terbaik.
6. Gunakan Optimal Pixel Adjustment Process pada citra untuk mengoptimalkan penyimpanan.
7. Hitung invers DCT 2D dari setiap blok.



3.2.2 Extraction Algorithm

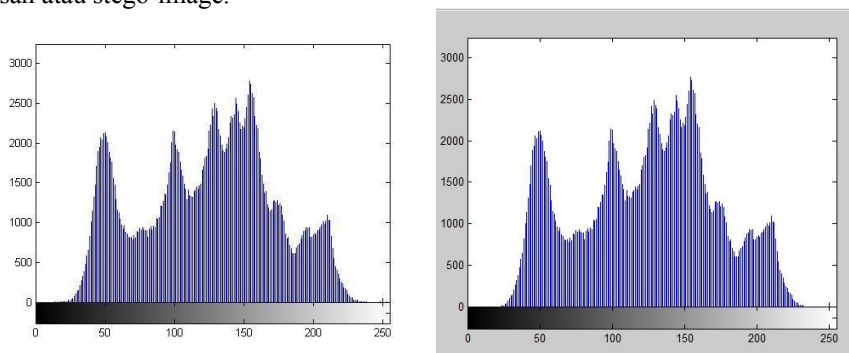
Berikut adalah diagram alir yang menjelaskan tentang proses ekstraksi pesan atau extraction process:

1. Bagi cover image kedalam blok 4x4.
2. Ekstrak koefisien transform domain dengan DCT 2D.
3. Gunakan fungsi yang didapatkan dalam fase penyisipan dan cari sekuensi pixel untuk ekstraksi.
4. Ekstrak k-LSBs pesan rahasia dalam setiap pixel.

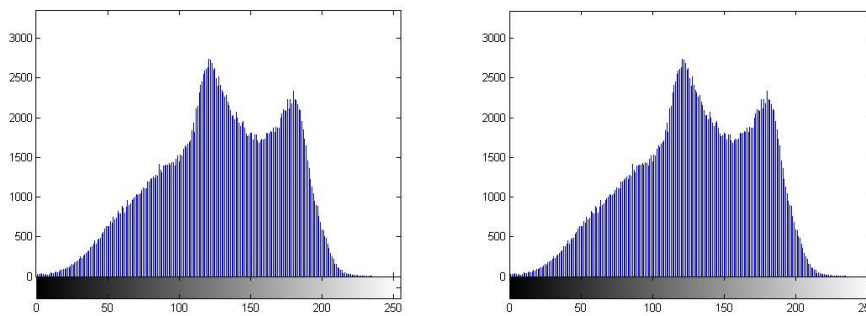


Sebuah pesan berupa teks yang akan dikirimkan kepada seseorang akan disembunyikan kedalam sebuah gambar/citra. Dalam hal ini teks dengan kalimat "Teks ini digunakan untuk pengujian steganography menggunakan teknik algoritma genetika yang akan disisipkan kedalam citra grayscale dan algoritma genetika yang digunakan dalam pemetaan. hasil psnr nya akan dibandingkan dengan teknik lain seperti dwt" akan disisipkan pada citra tersebut. Pada penelitian ini digunakan 4 buah citra (boat, jet, lena, dan baboon) gray-scale berukuran 512x512 dengan bit depth sebesar 8 bit.

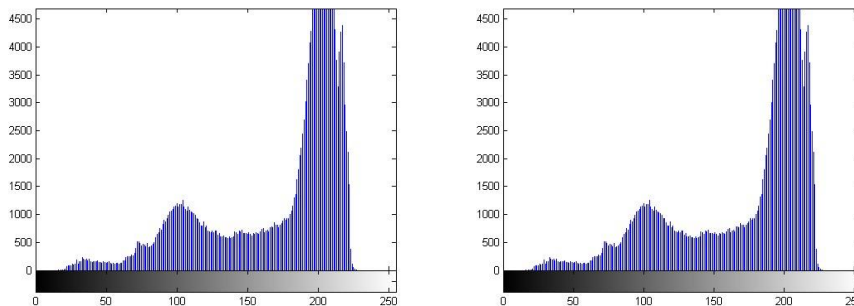
Algoritma genetika diterapkan saat mapping function dilakukan, evaluasi fitness dilakukan untuk memilih mapping function terbaik. Algoritma genetika menyisipkan n bit pesan kedalam koefisien dct secara random sehingga tidak ada perbedaan yang signifikan yang terlihat pada citra yang telah tersisipi pesan dengan citra asli. Hal ini bisa terlihat dari gambar 10,11,12 dan 13 histogram dari citra asli dengan histogram citra yang telah disisipi pesan atau stego-image.



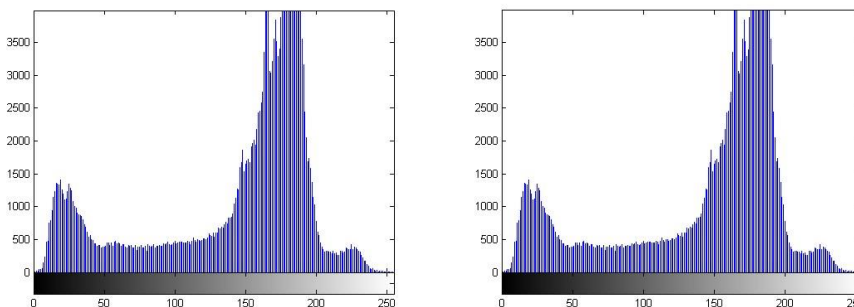
Gambar 10. Histogram Citra Lena Asli dan Histogram setelah proses steganography



Gambar 11. Histogram Citra Baboon Asli dan Histogram setelah proses steganography



Gambar 12. Histogram Citra Jet asli dan Histogram setelah proses steganography



Gambar 13. Histogram Citra Boat asli dan Histogram setelah proses steganography

Gambar diatas adalah citra stego-image yang telah disisipkan pesan dengan $k\text{-lsb}=3$. K-Least significant bit atau $k\text{-lsb}$ adalah bagian dari barisan data biner yang mempunyai nilai paling tidak berarti/paling kecil. Letaknya adalah paling kanan dari barisan bit. Histogram citra menggambarkan penyebaran nilai-nilai intensitas pixel dari suatu citra. Sumbu vertikal merupakan frekuensi/jumlah pixel sedangkan sumbu horizontal merupakan nilai intensitas warna. Dari histogram antara cover image dan stego image terlihat tidak ada perbedaan yang signifikan dari kedua citra tersebut setelah disisipkan pesan.

Perbandingan hasil dilakukan dengan melakukan proses steganografi dengan metode DCT dan algoritma genetika. Sedangkan metode yang dipakai dalam paper adalah metode DWT. Berikut adalah hasil dari metode yang digunakan dalam pengujian dan hasil penelitian se sebelumnya berdasarkan dengan paper [3], [5], [6], dan [7].

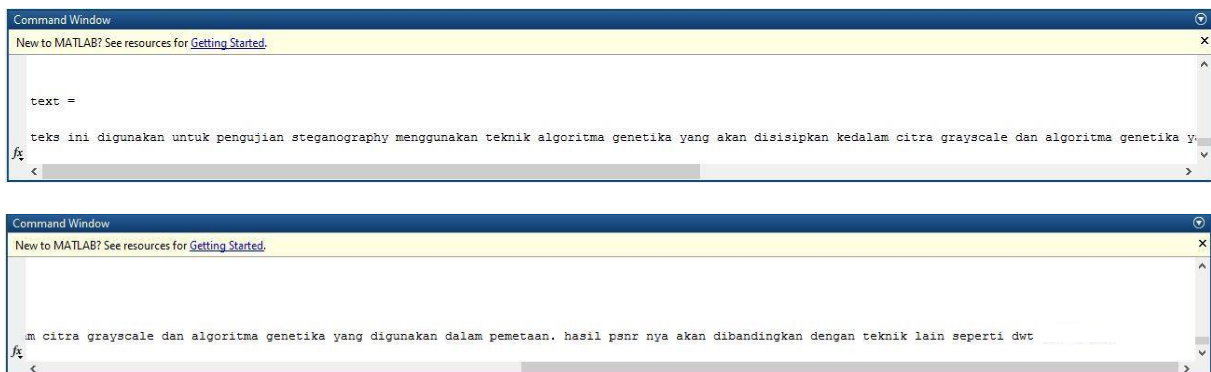
Tabel 1. perbandingan PSNR yang didapat dari hasil pengujian yang dilakukan dengan metode di paper [3], [5], [6], [7].

Image	Metode	PSNR	MSE
Boat	Metode yang diuji	60.23	0,061671
	GADWT [3]	39.94	
	DWT [6]	44.90	
	ADAPTIVE [5]	31.8	

	HDWT[7]	33.58	
Jet	Metode yang diuji	60.11	0,063399
	GADWT [3]	40.34	
	DWT [6]	44.96	
	ADAPTIVE [5]	30.89	
	HDWT[7]	32.69	
Lena	Metode yang diuji	66.79	0,013617
	GADWT [3]	45.20	
	DWT [6]	44.76	
	ADAPTIVE [5]	-	
	HDWT[7]	-	
Baboon	Metode yang diuji	61.88	0,042177
	GADWT [3]	40.44	
	DWT [6]	44.92	
	ADAPTIVE [5]	-	
	HDWT[7]	-	

Dari tabel 1 metode yang diujikan pada 4 image memiliki nilai PSNR lebih besar dibandingkan dengan 4 metode pada 4 referensi hal ini dikarenakan proses pemetaan pesan pada setiap pixel dalam citra menggunakan algoritma genetika secara random. Proses algoritma yang dilakukan secara random akan memperkecil peluang MSE. Semakin kecil MSE maka semakin besar pula nilai PSNR sehingga perbandingan kualitas citra sebelum dan sesudah proses steganography akan sama. Gambar akan mengalami sedikit perubahan ketika PSNR bernilai ≤ 35 , dan MSE bernilai ≥ 17 .

Teks yang terbaca atau yang diekstrak/dikeluarkan dari citra sama dengan teks yang dikirim dan dimasukkan kedalam citra sebelum proses steganography. Teks akan berubah ketika $k\text{-lsb} \geq 4$. Perubahan teks lebih bergantung kepada angka $k\text{-lsb}$, semakin kecil $k\text{-lsb}$ semakin mirip teks yang terbaca dengan teks yang dikeluarkan. Adapun teks yang disisipkan adalah kalimat “teks ini digunakan untuk pengujian steganography menggunakan teknik algoritma genetika yang akan disisipkan kedalam citra grayscale dan algoritma genetika yang digunakan dalam pemetaan. hasil psnr nya akan dibandingkan dengan teknik lain seperti dwt”. Hasil ekstraksi teks yang dimasukkan kedalam citra, dalam kasus ini adalah citra Lena, ditunjukkan pada gambar 13.



Gambar 14. Hasil ekstraksi teks yang dimasukkan kedalam citra Lena.

Pengujian juga dilakukan dengan citra yang mempunyai ukuran 256x256 untuk melihat toleransi. PSNR dan kapasitas penyimpanan citra tersebut dalam menyimpan teks yang sama.

Tabel 2. PSNR pada citra 256x256

Citra Ukuran 256x256	PSNR	MSE
Boat	53.81	0.27
Jet	54.42	0.23

Lena	53.80	0.27
Baboon	54.26	0.24

Hasil diatas menunjukkan bahwa PSNR pada citra dengan ukuran 256x256 yang sudah disisipi pesan masih tergolong baik namun lebih kecil daripada nilai PSNR pada citra dengan ukuran 512x512. Teks terbaca sempurna dengan $k\text{-lsb} = 3$.

KESIMPULAN

Algoritma Genetika berhasil diterapkan dalam menyembunyikan sebuah pesan kedalam citra dan berhasil mengoptimasi penyimpanan pesan dari hasil PSNR. PSNR yang dihasilkan pada penelitian ini sudah cukup baik mengingat nilai yang dihasilkan diatas 60 yang menunjukkan bahwa kualitas citra sebelum dan sesudah proses steganography tidak berubah. Dengan demikian pengiriman teks berhasil disembunyikan atau tidak dapat dicuri orang lain. Semakin besar ukuran citra maka semakin besar PSNR meskipun tidak terlalu signifikan.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada berbagai pihak yang telah membantu dalam penulisan makalah ini. Makalah ini didanai oleh Program Riset ITB 2018 dan Riset Pusat Unggulan Perguruan Tinggi (PUPT) kemenristekdikti 2018.

REFERENSI

1. N. Provos, P. Honeyman, "Hide and Seek: an introduction to steganography," IEEE Computer Society, pp. 32-44, Mei-Juni 2003
2. A.M. Fard, M.R Akbarzadeh and A. F Varasteh. "A New Genetic Algorithm Approach for Secure JPEG Steganography," International Conference on Engineering of Intelligence Systems, pp 1-6, 2006.
3. E. Ghaseni, J. Shanbehzadeh, N.Fassihi, "High Capacity Image Steganography using Wavelet Transform and Genetic Algorithm".
4. Suyanto, ST, MSc. 2008. Soft Computing Membangun Mesin Ber-IQ Tinggi. Bandung:informatika.
5. El Safy, R.O, Zayed. H. H, El Dessouki. A, "An adaptive steganography technique based on integer wavelet transform," ICNM International Conference on Networking and Media Convergence, pp 111-117, 2009.
6. P. Chen, H. Lin, "A DWT Based Approach for Image Steganography." International Journal of Applied Science and Engineering, Vol. 4, No. 3, pp. 275-290, 2006.
- 7.
8. B. Lai and L.Chang, "Adaptive Data Hiding for Images Based on Haar Discrete Wavelet transform," Lecture Notes in Computer Science, Vol 4319, 2006