

# Analisis Seleksi Tingkat Kecocokan Gambar pada MDID Multimedia Database Dengan Menggunakan Metode ImageDNA

## Analysis of Image Match Level Selection in MDID Multimedia Database Using the ImageDNA Method

Jimmy H. Moedjahedy<sup>1</sup>, Hamada Zein<sup>2</sup>, Isdayani B<sup>3</sup>,  
Erfan Tongalu<sup>4</sup>, Kusri<sup>5</sup>, M Syukri Mustafa<sup>6</sup>

Universitas Amikom Yogyakarta, Universitas Klabat Airmadidi  
Program Studi Magister Teknik Informatika Universitas Amikom Yogyakarta  
e-mail: \*jimmy@unklab.ac.id, hz831@umkt.ac.id, Isdayani.b@gmail.com,  
erfantongalu27@gmail.com, kusri@amikom.ac.id, syukri@dipanegara.ac.id

### Abstrak

Dengan semakin tersedianya pilihan informasi digital saat ini, definisi multimedia yang umum diterima adalah kombinasi dari berbagai media seperti teks, gambar, suara, video, animasi. Dalam teoritis basis data, multimedia basis data mulai dikenalkan yaitu kumpulan data multimedia terkait. Basis data yang dipilih untuk optimasi dalam penelitian ini adalah MDID (*Multiply Distorted Image Database*) yang terdiri dari 20 gambar referensi dan 1600 gambar yang sudah diberikan distorsi. Tujuan dari penelitian ini adalah untuk menguji kecocokan dari 1600 gambar tersebut dan dicocokkan dengan 20 gambar referensi dengan menggunakan metode *ImageDNA*. Nilai *ImageDNA* kemudian dilakukan uji data pencilaan, sehingga gambar yang nilai *ImageDNA*nya ekstrim akan dikeluarkan dari basis data MDID. Hasil dari penelitian ini adalah ada 100 gambar yang dikeluarkan dari basis data multimedia MDID.

**Kata kunci**— Multimedia Database, MDID, Image DNA

### Abstract

With the increasing availability of digital information choices today, the commonly accepted definition of multimedia is a combination of various media such as text, images, sound, video, animation. In database theorists, multimedia databases began to be introduced, namely related multimedia data sets. The database chosen for optimization in this study is MDID (*Multiply Distorted Image Database*) which consists of 20 reference images and 1600 images that have been given distortion. The purpose of this study was to test the suitability of 1600 images and match them with 20 reference images using the *ImageDNA* method. The *ImageDNA* value is then carried out with an outlier data test, so that images with extreme *ImageDNA* values will be removed from the MDID database. The results of this study are that there are 100 images that were removed from the MDID multimedia database.

**Keywords**— Multimedia Database, MDID, Image DNA

## 1. PENDAHULUAN

Basisdata multimedia adalah basisdata yang menyimpan data multimedia biasanya mencakup teks, gambar, audio, video, animasi, grafik, dan format kompleks lainnya. Dalam

beberapa tahun terakhir, pemrosesan data multimedia di komputer dan transmisi melalui jaringan berkembang pesat [1]. Dalam pencarian atau kueri informasi tekstual, ada beberapa metode yang biasa digunakan, namun metode-metode tersebut hanya untuk basisdata tradisional yang fokus kepada pencarian berbasis teks, sedangkan metode pencarian untuk gambar, suara atau video belum terlalu banyak dan masih dalam studi [2]. Beberapa penelitian terkait basisdata multimedia diantaranya MDID (Multiply Distorted Image Database) yang dilakukan oleh [3], basisdata wajah berdasarkan *text-video* [4], basisdata video multimodal biometric [5], basisdata UNMC-VIER yang berisi audio dan dan visual [6], basisdata GAID (Gait from Audio, Image and Depth) [7], basis data deteksi wajah KinectfaceDB [8]. Dari beberapa penelitian tersebut, yang akan menjadi referensi dengan tujuan optimasi adalah MDID. MDID merupakan basisdata multimedia yang terdiri dari 20 gambar referensi yang diambil dari beberapa basis data yang populer kemudian masing-masing gambar tersebut diberikan distorsi. Proses pemberian distorsi dengan menggunakan MATLAB, *segment linear function* dan *Kakadu tools*. Total gambar dalam basisdata tersebut adalah 1600 gambar yang didapatkan dari 20 gambar yang digandakan menjadi masing-masing 80 gambar yang terdistorsi. Adapun jenis distorsi yang digunakan adalah gaussian noise, gaussian blur, contrast change, JPEG, dan JPEG2000. Basisdata ini digunakan untuk mengukur kualitas metrik dari gambar. Tujuan dari penelitian ini adalah peneliti ingin menggunakan basisdata ini untuk menganalisa kecocokan gambar dari kumpulan dgambar yang ada. Selanjutnya, mengidentifikasi mana data yang sudah terlalu banyak distorsi sehingga gambar tersebut sudah berbeda jauh dengan gambar referensinya. Uji kecocokan diimplementasikan dengan menggunakan open source ImageDNA. Data observasi yang muncul dengan nilai ekstrim atau data pencilan akan diidentifikasi kemudian data tersebut akan dikeluarkan dari basisdata ini.

Adapun beberapa penelitian terkait dengan topik yang akan diteliti adalah yang dilakukan oleh [9] tujuan dari penelitian yang mereka lakukan adalah untuk mencocokkan properti gambar spasial yang menggunakan semua piksel yang diperoleh dalam gambar dengan cara, pertama, satu gambar dipilih sebagai referensi, sedangkan gambar lainnya atau disebut gambar sekunder dipindahkan sehubungan dengan yang pertama. Proses simetris yang dilakukan, tidak memperdulikan berapa banyak gambar yang dipilih sebagai referensi. Ada tiga parameter yang menentukan perpindahan gambar sekunder. Studi ini telah diuji dalam dua situasi representatif yang berbeda, yaitu sampel gambar dengan kontur yang jelas dan sampel gambar tanpa kontur yang ditentukan. Penelitian selanjutnya adalah yang dilakukan oleh [10], dimana tujuan dari penelitian mereka adalah untuk mencocokkan gambar berdasarkan template kecocokan gambar dengan menggunakan metode *fixed match pixel*. Hasil dari penelitian mereka adalah untuk gambar bebas noise, pencocokan gambar dapat diselesaikan secara akurat oleh pencocokan templat dasar tetapi kecepatan pencocokan gambar terlihat lambat. Sedangkan untuk gambar yang memiliki noise, pencocokan gambar dapat diselesaikan secara akurat oleh pencocokan templat dasar, tetapi kecepatan pencocokannya lambat. Berikutnya adalah studi yang dilakukan oleh [11], metode penelitian mereka adalah pertama, kecocokan awal diperoleh dengan menggunakan *ASIFT* dan kendala geometri epipolar. Kedua, pertandingan awal dibagi menjadi kelompok-kelompok yang berbeda berdasarkan kendala homografi dan matriks homografi diperkirakan menggunakan strategi berbasis *RANSAC* untuk setiap kelompok. Ketiga, fitur *FAST* dan batasan homografi digunakan untuk memperpanjang kecocokan awal. Akhirnya, kendala homografi diperbarui menggunakan pencocokan diperpanjang dan pencilan dihapus dengan kendala geometri epipolar. Penelitian yang dilakukan oleh [12] tujuan penelitian mereka adalah membangun deskriptor baru dengan menggabungkan deskriptor bentuk geometris yang diusulkan dan deskriptor SIFT tradisional untuk lebih meningkatkan tingkat pencocokan yang benar. Hasil eksperimen mereka menunjukkan validitas metode yang diusulkan. Kemudian penelitian yang dilakukan oleh [13], mereka merumuskan serangkaian ekstensi dari operator SIFT dan SURF, dengan mengganti pendeteksi titik bunga yang mendasari dan digunakan untuk menghitung deskriptor SIFT atau SURF oleh sekelompok titik minat ruang-skala umum.

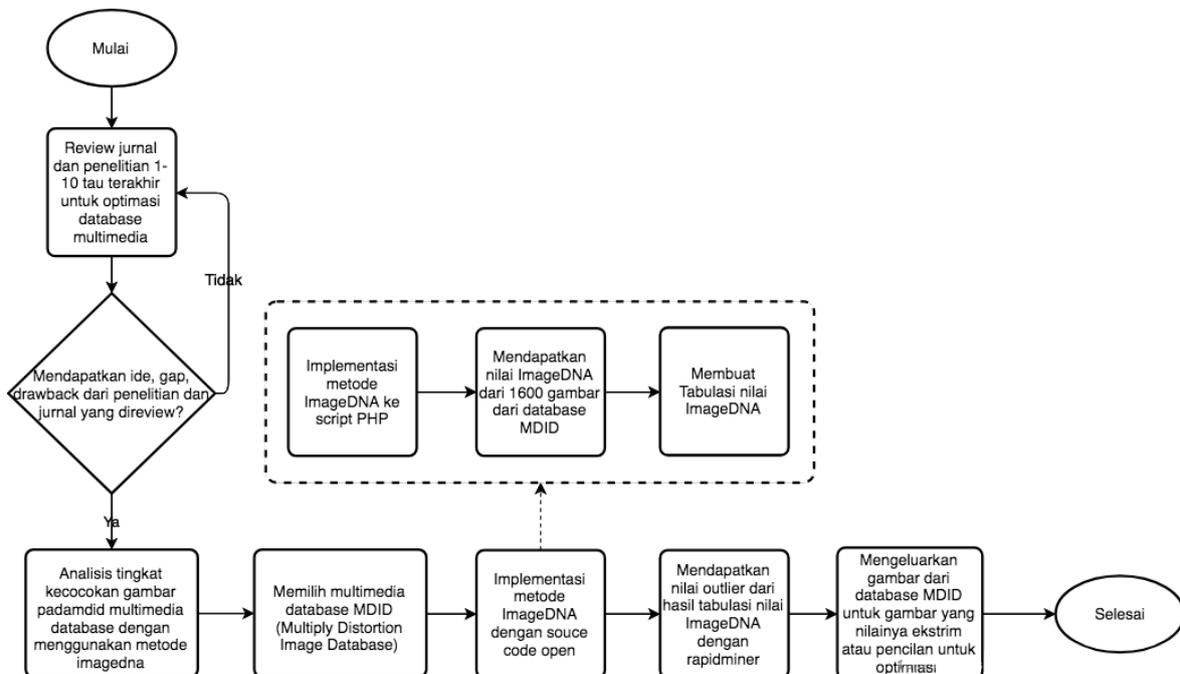
Kontribusi dari penelitian ini dapat dilihat pada tabel 1 yang merupakan rangkuman dari penelitian serupa berikut dengan perbedaannya dengan penelitian ini.

Tabel 1. Perbandingan dengan penelitian sebelumnya

Peneliti	Metode	Perbandingan
S. Piqueras Solsona, M. Maeder, R. Tauler, and A. de Juan [9]	Mencocokkan properti gambar dengan semua piksel	Peneliti menggunakan metode nilai ImageDNA
X. Xu and J. Zhang [10]	Mencocokkan gambar dengan metode <i>fixed match pixel</i>	Peneliti menggunakan metode nilai ImageDNA
Y. Zhai, G. Yu, H. Wang, and X. Guo [11]	Menghapus pencilan dengan menggunakan geometri epipolar	Peneliti menggunakan metode nilai ImageDNA untuk menghapus data pencilan
Q. Zhang, Y. Li, R. S. Blum, and P. Xiang [12]	Mencocokkan gambar dengan menggabungkan metode geometri SIFT tradisional	Peneliti menggunakan metode nilai ImageDNA
T. Lindeberg [13]	Mencocokkan gambar dengan metode SIFT dan SURF	Peneliti menggunakan metode nilai ImageDNA

## 2. METODE PENELITIAN

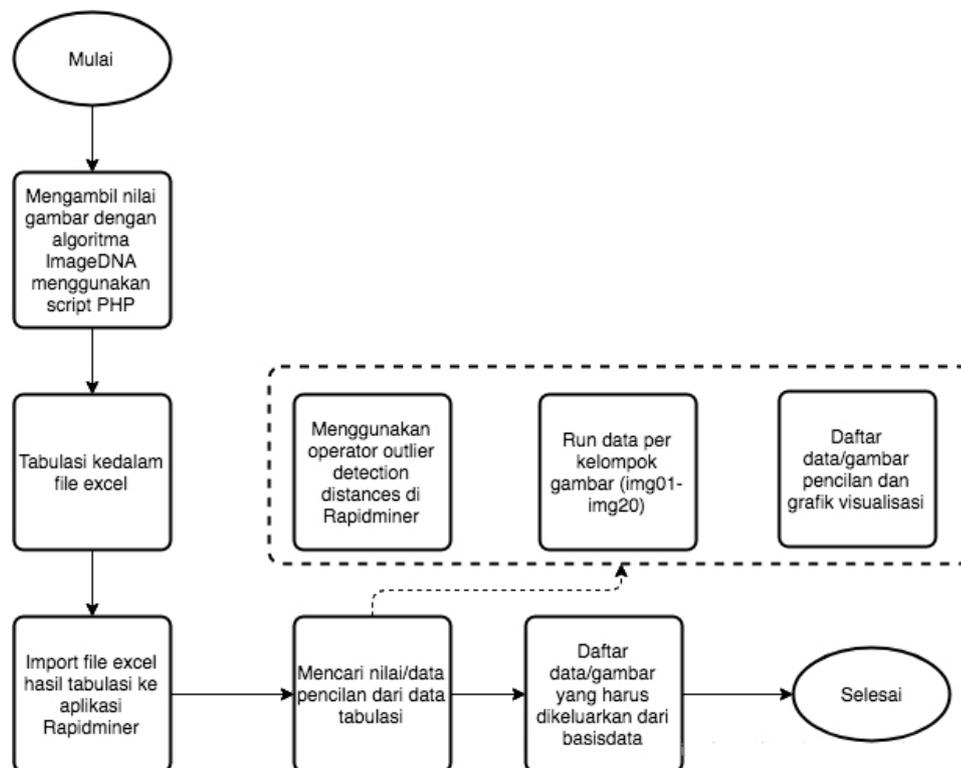
Jenis penelitian yang digunakan dalam penelitian ini adalah eksperimen, dimana peneliti menerapkan dan menggabungkan metode dalam hal ini ImageDNA. Selanjutnya sifat dari penelitian ini adalah deskriptif karena penulis akan mendeskripsikan mana gambar yang nilainya jauh berbeda dari gambar asli atau yang termasuk dalam data gambar pencilan. Pada gambar 1, merupakan diagram alir atau tahapan dalam menyelesaikan penelitian ini.



Gambar 1. Diagram Alur Penelitian

Tahapan penelitian yang dilakukan adalah:

1. Peneliti menentukan ide penelitian untuk topik optimasi multimedia database
2. Memilih MDID sebagai multimedia database yang terdiri dari 20 gambar referensi atau asli, masing-masing gambar telah diberikan distorsi sehingga menjadi 1600 gambar.
3. Masing-masing dari 1600 gambar tersebut akan dicari nilai ImageDNA dengan menerapkan script PHP open source yang diambil dari [14] Algoritma untuk mendapatkan DNA dari image adalah:
  - a. Konversi gambar ke hitam putih
  - b. Merubah ukuran gambar menjadi 512x512 piksel
  - c. Memecah-mecahkan gambar menjadi 32x32 piksel kemudian membuat rata-rata dari gambar piksel
  - d. Menggabungkan kembali gambar-gambar pada langkah c
  - e. Merubah ukuran gambar yang berukuran 1 piksel ke 16 piksel
  - f. Setelah semua langkah selesai, piksel dikonversi ke byte 0xFF dan disimpan dalam array yang bertindak sebagai DNA, yang menghasilkan 16x16 gambar DNA dalam 256 nilai. Ini dijumlahkan untuk menghitung delta yang mudah dilihat ketika membandingkan dengan DNA gambar lain.
4. Tabulasi hasil dari pencarian nilai ImageDNA
5. Menentukan nilai pencilan dari masing-masing gambar sesuai dengan gambar referensi dengan Rapidminer menggunakan *operator detect outlier distance*
6. Alur tabulasi nilai dan nilai pencilan dapat dilihat pada gambar 2



Gambar 2. Alur Tabulasi dan Pencarian Data Pencilan

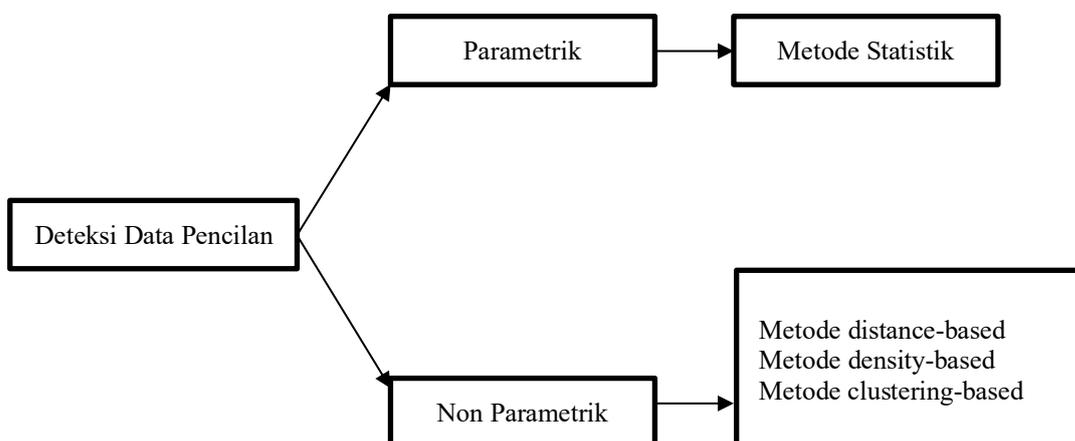
Pada gambar 1 dan 2, tanda dengan garis putus-putus adalah bagian lebih detail dari langkah-langkah yang ada Digambar.

### 2.1 Metode Pencarian Nilai PhotoDNA

ImageDNA mendapatkan ide untuk membuat program open source berdasarkan metode PhotoDNA yang dikeluarkan oleh Microsoft. Dalam website Microsoft [15] pada tahun 2009, Microsoft bermitra dengan Dartmouth College untuk mengembangkan PhotoDNA, sebuah teknologi yang membantu dalam menemukan dan menghapus gambar eksploitasi anak yang diketahui. Saat ini, PhotoDNA digunakan oleh organisasi di seluruh dunia dan telah membantu dalam pendeteksian, gangguan, dan pelaporan jutaan gambar eksploitasi anak. PhotoDNA membuat tanda tangan digital unik dikenal sebagai *hash* dari suatu gambar yang kemudian dibandingkan dengan tanda tangan atau *hashes* dari foto lain untuk menemukan salinan dari gambar yang sama. Ketika dicocokkan dengan database yang berisi hash dari gambar ilegal yang sebelumnya diidentifikasi, PhotoDNA adalah alat yang luar biasa untuk membantu mendeteksi, mengganggu dan melaporkan distribusi bahan eksploitasi anak. PhotoDNA bukan perangkat lunak pengenalan wajah dan tidak dapat digunakan untuk mengidentifikasi seseorang atau objek dalam suatu gambar. Hash PhotoDNA tidak dapat dibalik, dan oleh karena itu tidak dapat digunakan untuk membuat ulang gambar

### 2.2 Metode Pencarian Data Pencilan

Menurut [16] dalam buku mereka, sebuah obyek dalam data disebut data pencilan jika menyimpang dari perilaku normal data tersebut, nilai yang jauh dari nilai yang diharapkan, atau tidak mirip dengan objek lain dalam hal karakteristiknya. Ada 3 karakteristik dalam data pencilan yaitu berdasarkan ukuran, berdasarkan keragaman dan berdasarkan peran. Ada 2 taksonomi dalam pendeteksian data pencilan, yaitu parametrik dan non parametrik seperti yang ditunjukkan pada gambar 3. Jika masuk dalam kategori parametrik maka metode statistik yang akan digunakan, sedangkan jika data masuk dalam kategori non parametrik maka bisa menggunakan 3 pendekatan yaitu metode *distance-based*, metode *density-based*, dan metode *clustering-based*.



Gambar 3. Taksonomi Deteksi data Pencilan

Ada beberapa metode pencarian data pencilan, dalam [17] metode pertama adalah *average k-nearest neighbors detection*, yang adalah varian yang kuat dari detektor *k-nearest neighbor* yang tepat. Gagasan dasar dalam metode ini adalah untuk menghitung jarak rata-rata titik data ke *k-nearest neighbors* sebagai nilai pencilan. Kemudian metode yang berikutnya adalah

*harmony k-nearest neighbor detector*, yang merupakan metode yang jarang digunakan, tetapi merupakan teknik dengan beberapa sifat menarik dan tidak biasa. Seperti *average k-nearest neighbors detection*, ide dasarnya adalah menggunakan rata-rata dari semua jarak k ke tetangga terdekat k; perbedaan utamanya adalah bahwa rata-rata harmonik digunakan sebagai ganti rata-rata aritmatika. Berikutnya adalah Local Outlier Factor (LOF), dimana LOF menormalkan jarak tetangga k-terdekat dengan jarak di lokalitasnya. Misalnya, jika suatu wilayah diharapkan mengandung nilai yang lebih kecil dari jarak tetangga k-terdekat, maka skor tersebut disesuaikan dengan fakta ini. Ide dasarnya adalah untuk menyesuaikan skor pencilan untuk variasi dalam kepadatan di berbagai lokasi data.

### 3. HASIL DAN PEMBAHASAN

Setelah dilakukan pencarian semua nilai ImageDNA dari 1600 gambar, nilai ImageDNA seperti terlihat pada gambar 4 yaitu nilai sesudah tanda sama dengan. Nilai ini didapat dengan menjalankan script PHP yang sudah dimodifikasi untuk basisdata MDID. Langkah selanjutnya adalah tabulasi nilai-nilai tersebut.

```

localhost/imageDNA/test20.php

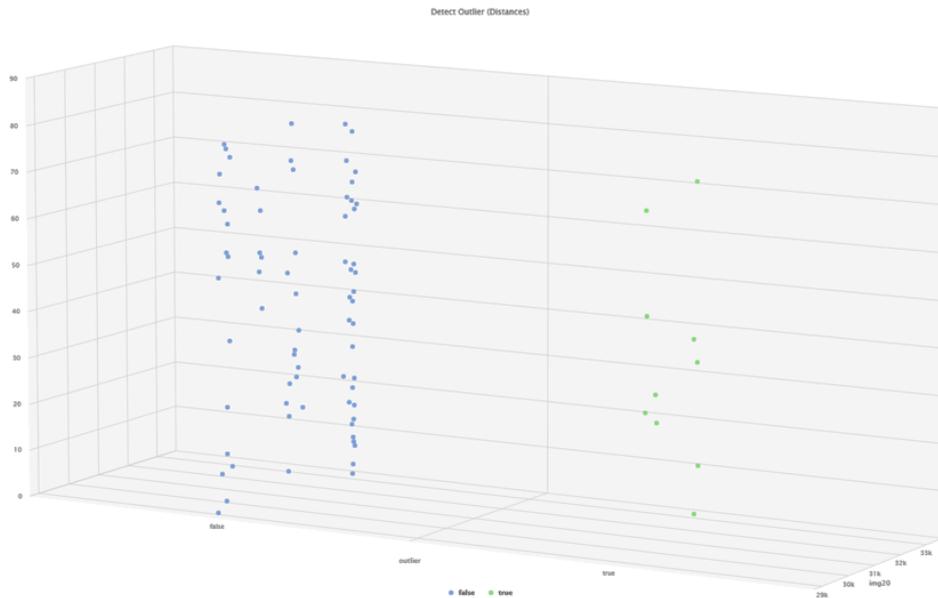
ImageDNA:
106,28,47,27,42,111,224,233,229,240,226,197,194,195,208,218,147,138,92,25,32,37,147,222,23
1,235,208,202,212,198,189,189,143,191,211,151,122,140,196,227,214,206,195,211,218,221,195,
191,161,171,221,217,221,220,224,237,218,186,184,185,200,176,174,205,141,189,193,222,229,19
9,74,96,139,181,168,146,147,141,173,214,107,189,185,193,200,116,14,22,53,107,153,142,106,1
33,200,228,140,103,110,83,60,24,28,28,26,53,67,67,90,162,179,184,185,53,12,7,7,21,23,27,28
,51,57,91,94,169,197,170,180,104,20,13,33,34,23,25,55,65,102,135,162,168,179,189,173,171,1
59,98,86,100,143,108,118,147,186,167,157,173,182,212,123,115,88,43,53,95,148,125,119,154,1
91,198,166,178,197,228,70,16,2,5,65,52,104,128,157,144,171,169,185,203,196,214,19,4,4,2,13
,94,108,135,177,165,150,168,209,227,227,212,25,3,5,5,34,67,140,134,148,147,167,168,193,224
,227,214,23,3,3,3,11,45,125,147,137,161,200,204,182,163,144,148,30,4,2,5,12,43,65,126,162,
159,192,215,153,83,28,23 = 33442
ImageDNA:
103,35,57,33,46,96,198,203,196,217,192,143,139,140,160,171,130,121,86,32,38,46,134,188,206
,216,162,151,170,149,138,138,128,148,172,130,109,132,169,196,171,158,141,166,180,186,146,1
38,133,137,187,178,185,184,191,219,181,139,141,142,149,142,136,155,127,144,144,189,203,153
,69,90,116,138,136,133,132,132,139,169,111,149,140,147,153,96,19,27,58,96,124,130,109,121,
158,191,117,105,92,73,57,31,35,34,32,62,78,80,90,130,139,142,139,52,15,10,11,25,29,33,34,5
6,65,97,89,136,151,135,135,89,25,18,37,41,28,32,61,72,99,123,129,132,140,138,135,136,132,1
07,94,99,125,120,124,132,140,136,133,136,139,163,125,125,101,53,64,102,131,126,125,133,143
,149,136,139,152,192,81,22,6,8,78,64,122,126,134,131,135,136,143,163,143,166,21,6,6,5,18,1
14,124,129,138,136,133,136,165,197,200,164,30,6,9,8,42,83,129,129,132,132,136,135,151,191,
200,170,28,7,6,6,16,56,123,132,129,133,146,156,142,135,128,124,36,6,5,8,17,50,76,126,135,1
34,149,171,129,87,35,27 = 28757
ImageDNA:
105,28,52,29,43,108,218,226,225,238,221,195,188,191,202,211,144,140,92,29,35,36,146,221,22
9,235,206,203,212,197,188,186,140,192,209,151,121,137,195,224,209,204,191,211,217,220,196,
186,161,168,220,216,223,219,221,234,215,187,186,185,202,178,171,201,139,190,197,220,226,19
6,74,95,140,183,171,142,146,142,175,208,108,189,189,191,201,118,15,24,54,109,158,140,106,1
35,197,219,137,100,110,83,60,24,31,28,29,54,65,66,91,166,180,182,179,52,12,12,12,23,22,28,
33,52,57,91,93,170,195,168,172,101,23,15,35,37,24,26,56,67,102,135,166,168,181,190,169,170
,158,98,87,100,142,108,118,147,186,169,161,177,182,207,120,117,90,45,55,95,145,121,117,152
,191,199,173,182,195,221,69,22,6,9,69,52,102,127,155,142,172,169,188,201,195,210,23,10,11,
3,16,96,105,134,182,167,148,172,207,224,225,207,30,9,12,10,37,66,138,133,149,145,169,171,1
92,222,224,211,27,8,7,6,16,47,126,147,133,165,200,202,182,165,148,147,33,6,3,11,17,46,68,1
28,165,162,190,213,153,87,28,25 = 33463

```

Gambar 4. Hasil Pencarian Nilai ImageDNA

Setelah 1600 gambar ditambah dengan 20 gambar referensi didapatkan nilai ImageDNA, maka masing-masing nilai tersebut pada setiap kategori gambar dicari data pencilan. Data

pencilan dicari dengan menggunakan Rapidminer *operator outlier detection distances* dengan pengaturan 10 data outlier yang akan dikeluarkan pada 20 kelompok gambar. Pada gambar 5 merupakan salah satu hasil grafik visualisasi, titik berwarna hijau merupakan data pencilan sehingga pada tabel 1 dan tabel 2 merupakan hasil tabulasi gambar yang harus dikeluarkan dari database MDID.



Gambar 5. Grafik Hasil Data Pencilan

Tabel 2. Hasil deteksi data pencilan pada img01-img10

img01	img02	img03	img04	img05	img06	img07	img08	img09	img10
img01_09	img02_03	img03_05	img04_11	img05_14	img06_01	img07_16	img08_21	img09_05	img10_06
img01_17	img02_09	img03_10	img04_19	img05_25	img06_06	img07_21	img08_22	img09_09	img10_09
img01_21	img02_14	img03_15	img04_20	img05_27	img06_09	img07_23	img08_24	img09_10	img10_26
img01_37	img02_18	img03_19	img04_30	img05_31	img06_14	img07_30	img08_28	img09_14	img10_30
img01_44	img02_21	img03_40	img04_36	img05_32	img06_32	img07_33	img08_35	img09_16	img10_39
img01_48	img02_37	img03_52	img04_50	img05_34	img06_43	img07_46	img08_47	img09_22	img10_40
img01_50	img02_52	img03_57	img04_52	img05_54	img06_62	img07_56	img08_51	img09_26	img10_42
img01_55	img02_66	img03_60	img04_64	img05_58	img06_68	img07_60	img08_61	img09_37	img10_51
img01_56	img02_72	img03_63	img04_66	img05_69	img06_76	img07_66	img08_64	img09_57	img10_53
img01_62	img02_75	img03_78	img04_69	img05_70	img06_78	img07_73	img08_65	img09_72	img10_77

Tabel 2 merupakan data atau gambar dari gambar 1 sampai gambar 10 yang merupakan hasil tabulasi dari aplikasi Rapidminer. Untuk kelompok gambar 1, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 01\_09, 01\_17, 01\_21, 01\_37, 01\_44, 01\_48, 01\_50, 01\_55, 01\_56, 01\_62. Untuk kelompok gambar 2, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 02\_03, 02\_09, 02\_14, 02\_18,

02\_21, 02\_37, 02\_52, 02\_66, 02\_72, 02\_75. Untuk kelompok gambar 3, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 03\_05, 03\_10, 03\_15, 03\_19, 03\_40, 03\_52, 03\_57, 03\_60, 03\_63, 03\_78. Untuk kelompok gambar 4, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 04\_11, 04\_19, 04\_20, 04\_30, 04\_36, 04\_50, 04\_52, 04\_64, 04\_66, 04\_69. Untuk kelompok gambar 5, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 05\_14, 05\_25, 05\_27, 05\_31, 05\_32, 05\_34, 05\_54, 05\_58, 05\_69, 05\_70. Untuk kelompok gambar 6, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 06\_01, 06\_06, 06\_09, 06\_14, 06\_32, 06\_43, 06\_62, 06\_68, 06\_76, 06\_78. Untuk kelompok gambar 7, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 07\_16, 07\_21, 07\_23, 07\_30, 07\_33, 07\_46, 07\_56, 07\_60, 07\_66, 07\_73. Untuk kelompok gambar 8, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 08\_21, 08\_22, 08\_24, 08\_28, 08\_35, 08\_47, 08\_51, 08\_61, 08\_64, 08\_65. Untuk kelompok gambar 9, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 09\_05, 09\_09, 09\_10, 09\_14, 09\_16, 09\_22, 09\_26, 09\_37, 09\_57, 09\_72. Untuk kelompok gambar 10, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 10\_06, 10\_09, 10\_26, 10\_30, 10\_39, 10\_40, 10\_42, 10\_51, 10\_53, 10\_77.

*Tabel 3. Hasil deteksi data pencilan pada img11-img20*

img11	img12	img13	img14	img15	img16	img17	img18	img19	img20
img11_02	img12_05	img13_01	img14_01	img15_16	img16_19	img17_03	img18_07	img19_10	img20_04
img11_06	img12_14	img13_03	img14_16	img15_17	img16_20	img17_10	img18_10	img19_16	img20_14
img11_17	img12_19	img13_26	img14_26	img15_21	img16_28	img17_22	img18_30	img19_20	img20_26
img11_20	img12_24	img13_30	img14_29	img15_32	img16_33	img17_23	img18_31	img19_28	img20_29
img11_42	img12_31	img13_58	img14_34	img15_37	img16_35	img17_40	img18_32	img19_36	img20_32
img11_47	img12_44	img13_61	img14_50	img15_44	img16_50	img17_43	img18_62	img19_52	img20_36
img11_57	img12_67	img13_62	img14_61	img15_50	img16_70	img17_51	img18_68	img19_53	img20_41
img11_67	img12_68	img13_66	img14_63	img15_56	img16_71	img17_60	img18_71	img19_54	img20_49
img11_72	img12_69	img13_67	img14_66	img15_62	img16_75	img17_73	img18_74	img19_62	img20_71
img11_74	img12_71	img13_79	img14_76	img15_78	img16_79	img17_74	img18_80	img19_79	img20_74

Tabel 3 merupakan data atau gambar dari gambar 11 sampai gambar 20 yang merupakan hasil tabulasi dari aplikasi Rapidminer. Untuk kelompok gambar 11, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 11\_02, 01\_16, 11\_17, 11\_20, 11\_42, 11\_47, 11\_57, 11\_67, 11\_72, 11\_74. Untuk kelompok gambar 12, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 12\_05, 12\_14, 12\_19, 12\_24, 12\_31, 12\_44, 12\_67, 12\_68, 12\_69, 12\_71. Untuk kelompok gambar 13, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 13\_01, 13\_03, 13\_26, 13\_30, 13\_58, 13\_61, 13\_62, 13\_66, 13\_67, 13\_79. Untuk kelompok gambar 14, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 14\_01, 14\_16, 14\_26, 14\_29, 14\_34, 14\_50, 14\_61, 14\_63, 14\_66, 14\_76. Untuk kelompok gambar 15, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 15\_16, 15\_17, 15\_21, 15\_32, 15\_37, 15\_44, 15\_50, 15\_56, 15\_62, 15\_78. Untuk kelompok gambar 16, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 16\_19, 16\_20, 16\_28, 16\_33, 16\_35, 16\_50, 16\_70, 16\_71, 16\_75, 16\_79. Untuk kelompok gambar 17, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 17\_03, 17\_10, 17\_22, 17\_23, 17\_40, 17\_43, 17\_51, 17\_60, 17\_73, 17\_74. Untuk kelompok gambar 18, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 18\_07, 18\_10, 18\_30, 18\_31, 18\_32, 18\_62, 18\_68, 18\_71, 18\_74, 18\_80. Untuk kelompok gambar 19, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 19\_10, 19\_16, 19\_20, 19\_28,

19\_36, 19\_52, 19\_53, 19\_54, 19\_62, 19\_79. Terakhir, untuk kelompok gambar 20, dari 80 gambar yang terdistorsi, ada 10 gambar yang akan dikeluarkan yaitu gambar 20\_04, 20\_14, 20\_26, 20\_29, 20\_32, 20\_36, 20\_41, 20\_49, 20\_71, 20\_74.

#### 4. KESIMPULAN

Setelah mendapatkan hasil penelitian, kesimpulan dari penelitian ini adalah agar terjadi peningkatan waktu pencarian dan tingkat kecocokan data dengan menggunakan basisdata MDID maka 100 gambar dari 1600 gambar yang didistorsi harus dikeluarkan sebelum menguji menggunakan algoritma kecocokan gambar. Hal ini disebabkan, 100 gambar yang dikeluarkan tersebut, nilai ImageDNAnya sudah melewati gambar referensi yang diambil. Adapun gambar-gambar yang harus dikeluarkan dapat dilihat pada tabel 2 dan tabel 3.

#### 5. SARAN

Saran dari penelitian ini adalah, bagi para peneliti lainnya yang ingin mengembangkan penelitian ini, dapat melakukan uji coba dengan implementasi algoritma tingkat kecocokan gambar. Kemudian membandingkan waktu pencarian dengan menghilangkan 100 gambar seperti pada tabel 1 dan tabel 2 kemudian menguji kembali menggunakan semua gambar yang ada pada basisdata MDID.

#### UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ibu Dr. Kusri, M. Kom dan Bapak M. Syukri Mustafa, S.Si., M.MSI yang telah memberi arahan dan bantuan dalam menyelesaikan penelitian ini.

#### DAFTAR PUSTAKA

- [1] X. Gao, Y. Zhu, D. Kim, J. Li, and W. Wu, "A novel multi-channel data broadcast scheme for MultiMedia DataBase Systems," in *Proceedings of the International Conference on Parallel and Distributed Systems - ICPADS*, 2012, pp. 132–139.
- [2] P. K. Yadav and S. Rizvi, "An exhaustive study on data mining techniques in mining of Multimedia database," in *Proceedings of the 2014 International Conference on Issues and Challenges in Intelligent Computing Techniques, ICICT 2014*, 2014, pp. 541–545.
- [3] W. Sun, F. Zhou, and Q. Liao, "MDID: A multiply distorted image database for image quality assessment," *Pattern Recognition*, pp. 153–168, 2017.
- [4] D. Zeng, Y. Bao, K. Liu, F. Zhao, and Q. Tian, "Face database generation based on text–video correlation," *Neurocomputing*, 2016.
- [5] B. Ríos-Sánchez *et al.*, "Gb2μMOD: A MUltiMODal biometric video database using visible and IR light," *Information Fusion*, 2016.
- [6] Y. W. Wong *et al.*, "A new multi-purpose audio-visual UNMC-VIER database with multiple variabilities," *Pattern Recognition Letters*, vol. 31, pp. 1503–1510, 2011.
- [7] M. Hofmann, J. Geiger, S. Bachmann, B. Schuller, and G. Rigoll, "The TUM Gait from Audio, Image and Depth (GAID) database: Multimodal recognition of subjects and traits," *Journal of Visual Communication and Image Representation*, vol. 5, pp. 195–206, 2014.

- 
- [8] R. Min, N. Kose, and J. L. Dugelay, "KinectfaceDB: A kinect database for face recognition," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2014.
- [9] S. Piqueras Solsona, M. Maeder, R. Tauler, and A. de Juan, "A new matching image preprocessing for image data fusion," *Chemometrics and Intelligent Laboratory Systems*, no. 164, pp. 32–42, 2017.
- [10] X. Xu and J. Zhang, "The method of image matching by taking every fixed match pixel," in *Proceedings - 2012 5th International Symposium on Computational Intelligence and Design, ISCID 2012*, 2012.
- [11] Y. Zhai, G. Yu, H. Wang, and X. Guo, "Image matching for structured scenes based on ASIFT and homography constraint," in *2017 3rd IEEE International Conference on Computer and Communications, ICC3 2017*, 2018.
- [12] Q. Zhang, Y. Li, R. S. Blum, and P. Xiang, "Matching of images with projective distortion using transform invariant low-rank textures," *Journal of Visual Communication and Image Representation*, vol. 38, pp. 602–613, 2016.
- [13] T. Lindeberg, "Image Matching Using Generalized Scale-Space Interest Points," *Journal of Mathematical Imaging and Vision*, vol. 52, pp. 3–36, 2015.
- [14] Estella, "ImageDNA," 2019. [Online]. Available: <https://github.com/Estella/ImageDNA>. [Accessed: 12-Jan-2020].
- [15] Microsoft, "PhotoDNA," 2009. [Online]. Available: <https://www.microsoft.com/en-us/photodna>. [Accessed: 12-Jan-2020].
- [16] N. N. R. R. S. G. A. N. M. M, *Outlier Detection: Techniques and Applications: A Data Mining Perspective*. Gewerbestrasse: Springer Nature Switzerland, 2019.
- [17] C. C. Aggarwal and S. Sathe, *Outlier ensembles: An introduction*. 2017.