

**Enhancing Texture-Based Image Retrieval using GLCM and
DBSCAN on a Multifaceted Dataset**

Laporan Tugas Akhir

Diajukan Untuk Memenuhi
Persyaratan Guna Meraih Gelar Sarjana
Informatika Universitas Muhammadiyah Malang



Muhammad Rivaldi Asyhari
201910370311370

Bidang Minat
Data Science

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MALANG
2023**

LEMBAR PERSETUJUAN

TWITTER SENTIMEN ANALISIS DENGAN GAUSSIAN NAIVE BAYES

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang

Menyetujui,

Malang, 29 September 2023

Dosen Pembimbing 1



Vinna Rahmayanti S.Si., M.Si

NIP. 180306071990PNS.

Dosen Pembimbing 2



Ir. Yufis Azhar S.Kom., M.Kom.

NIP. 10814100544PNS.

LEMBAR PENGESAHAN
TWITTER SENTIMEN ANALISIS DENGAN GAUSSIAN
NAIVE BAYES
TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

MUHAMMAD RIVALDI ASYHARI

201910370311370

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji
pada tanggal 29 September 2023

Menyetujui,

Dosen Penguji 1



Wildan Suharso S.Kom., M.Kom

NIP. 10817030596PNS.

Dosen Penguji 2



Briansyah Setio Wiyono S.Kom.,

M.Kom

NIP. 190913071987PNS.

Mengetahui,
Ketua Jurusan Informatika



Wicaksono S.kom. M.Cs.
NIP. 10814100541PNS.

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

NAMA : MUHAMMAD RIVALDI ASYHARI

NIM : 201910370311370

FAK./JUR. : Informatika

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul **“TWITTER SENTIMEN ANALISIS DENGAN GAUSSIAN NAIVE BAYES”** beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,
Dosen Pembimbing



Vinna Rahmayanti S S.Si., M.Si

Malang, 29 September 2023
Yang Membuat Pernyataan



MUHAMMAD RIVALDI
ASYHARI

ABSTRAK

Pengambilan gambar berbasis tekstur merupakan aspek penting dari berbagai aplikasi visi komputer. Dalam penelitian kami, kami telah mengusulkan metode untuk meningkatkan pengambilan gambar berbasis tekstur dengan memanfaatkan ekstraksi fitur Gray Level Co-occurrence Matrix (GLCM) dan pengelompokan Cluster Spasial Berbasis Kepadatan Aplikasi dengan Noise (DBSCAN). Untuk mengevaluasi pendekatan kami, kami telah menggunakan dataset Corel 10k, yang terdiri dari 10.000 gambar beragam dari berbagai kategori. Metodologi kami melibatkan beberapa langkah. Pertama, kami mengonversi gambar menjadi skala abu-abu dan menormalkan nilai piksel. Kemudian kami mengekstrak empat fitur penting dari GLCM: entropi, energi, kontras, dan korelasi. Fitur-fitur ini memainkan peran kunci dalam menentukan kesamaan gambar. Selanjutnya, kami menerapkan pengelompokan DBSCAN untuk menyempurnakan hasil pengambilan berdasarkan fitur GLCM ini. Untuk menilai kinerja pendekatan kami, kami menggunakan metrik jarak yang berbeda seperti Euclidean, City Block, Bray Curtis, dan Canberra. Melalui analisis eksperimental, kami telah memperoleh hasil yang menjanjikan yang menyoroti efektivitas metode yang kami usulkan. Pendekatan GLCM DBSCAN secara konsisten mengungguli penggunaan GLCM saja dalam hal presisi pengambilan. Di antara metrik jarak yang digunakan untuk evaluasi, jarak Canberra mencapai nilai presisi tertinggi untuk mengukur kesamaan antara fitur berbasis GLCM dalam himpunan data Corel 10k. Ini menunjukkan kesesuaiannya sebagai ukuran kesamaan dalam konteks ini. Secara keseluruhan, penelitian kami berkontribusi untuk meningkatkan pengambilan gambar berbasis tekstur dengan menggunakan ekstraksi fitur GLCM dan metode pengelompokan DBSCAN. Hasil evaluasi yang berhasil memvalidasi efektivitas pendekatan kami dan menawarkan wawasan berharga untuk perbaikan di masa depan di bidang ini.

Kata Kunci: Corel-10k, GLCM, DBSCAN

ABSTRACT

Texture based image retrieval is an important aspect of various computer vision applications. In our research, we have proposed a method to enhance texture based image retrieval by utilizing Gray Level Co-occurrence Matrix (GLCM) feature extraction and Density Based Spatial Clustering of Applications with Noise (DBSCAN) clustering. To evaluate our approach, we have utilized the Corel 10k dataset, which consists of 10,000 diverse images from different categories. Our methodology involves several steps. Firstly, we convert the images to grayscale and normalize the pixel values. Then we extract four significant features from the GLCM: entropy, energy, contrast, and correlation. These features play a key role in determining image similarity. Subsequently, we apply DBSCAN clustering to refine the retrieval results based on these GLCM features. To assess the performance of our approach, we employ different distance metrics such as Euclidean, City Block, Bray Curtis and Canberra. Through experimental analysis, we have obtained promising results that highlight the effectiveness of our proposed method. The GLCM DBSCAN approach consistently outperforms using GLCM alone when it comes to retrieval precision. Among the distance metrics used for evaluation, Canberra distance achieves the highest precision values for measuring similarity between GLCM based features in the Corel 10k dataset. This indicates its suitability as a measure of similarity in this context. Overall, our research contributes to enhancing texture based image retrieval by employing GLCM feature extraction and DBSCAN clustering methods. The successful evaluation results validate the effectiveness of our approach and offer valuable insights for future improvements in this field.

Keywords: *Corel-10k, GLCM, DBSCAN*

LEMBAR PERSEMBAHAN

Puji syukur kepada Allah SWT atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan Tugas Akhir ini. Penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Yufis Azhar, S.Kom., M.Kom. dan Vinna Rahmayanti Setyaning Nastiti, S.Si., M.Si. selaku pembimbing tugas akhir.
2. Bapak/Ibu Dekan Fakultas Teknik Universitas Muhammadiyah Malang.
3. Bapak Galih Wasis Wicaksono, S.Kom., M.Cs. selaku Ketua Program Studi Informatika Universitas Muhammadiyah Malang.
4. Kedua orang tua saya, Bapak Sumardi (Almarhum) dan Ibu R. Franti Astuti yang selalu memberi do'a dan dukungan sehingga penulis dapat sampai pada titik ini.
5. Teman saya Yusril yang telah memperbolehkan saya menginap di kosnya.
6. Motor saya yang sudah menemani saya selama perjalanan pergi maupun pulang tanpa masalah.
7. Terimakasih untuk diri saya sendiri yang telah berusaha dan mampu bertahan hingga sampai pada titik ini sehingga dapat menyelesaikan tugas akhir ini dengan baik.

Malang, 29 September 2023



Digndal dengan CamScanner

Muhammad Rivaldi Asyhari

KATA PENGANTAR

Dengan memanjatkan puji syukur kehadirat Allah SWT. Atas limpahan rahmat dan hidayah-NYA sehingga peneliti dapat menyelesaikan tugas akhir yang berjudul

“Enhancing Texture-Based Image Retrieval using GLCM and DBSCAN on a Multifaceted Dataset”

Pada penelitian ini disajikan pokok-pokok bahasan yang meliputi latar belakang, metode penelitian, serta hasil pembahasan dari proses penelitian ini dan telah disimpulkan berdasarkan hasil yang telah didapatkan pada proses penelitian ini.

Peneliti menyadari sepenuhnya bahwa dalam penulisan tugas akhir ini masih banyak kekurangan dan keterbatasan. Oleh karena itu peneliti mengharapkan saran yang membangun agar tulisan ini bermanfaat bagi perkembangan ilmu pengetahuan.

Malang, 29 September 2023



CS Digiindal dengan CamScanner

Muhammad Rivaldi Asyhari

DAFTAR ISI

HALAMAN JUDUL	
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
BAB I	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan Penelitian	2
1.4 Cakupan Masalah	2
BAB II	3
2.1 Penelitian Terdahulu	3
2.2 GLCM	3
2.3 DBSCAN	3
BAB III	4
3.1 Corel-10k	4
3.2 Ekstraksi Fitur menggunakan GLCM	4
3.3 Pengelompokan DBSCAN	6
3.4 <i>Evaluation Metrics</i>	7
BAB IV	8
BAB V	10
DAFTAR PUSTAKA	11

DAFTAR PUSTAKA

- [1]. S. R. Dubey, "A Decade Survey of Content Based Image Retrieval Using Deep Learning," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 5, pp. 2687–2704, May 2022, doi: 10.1109/TCSVT.2021.3080920.
- [2]. Y. Li, J. Ma, and Y. Zhang, "Image retrieval from remote sensing big data: A survey," *Information Fusion*, vol. 67, pp. 94–115, Mar. 2021, doi: 10.1016/j.inffus.2020.10.008.
- [3]. X. Li, J. Yang, and J. Ma, "Recent developments of content-based image retrieval (CBIR)," *Neurocomputing*, vol. 452, pp. 675–689, Sep. 2021, doi: 10.1016/j.neucom.2020.07.139.
- [4]. M. K. Alsmadi, "Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features," *Arab J Sci Eng*, vol. 45, no. 4, pp. 3317–3330, Apr. 2020, doi: 10.1007/s13369-020-04384-y.
- [5]. S. Lee, H. Seong, S. Lee, and E. Kim, "Correlation Verification for Image Retrieval." [Online]. Available: <https://github.com/sungonce/CVNet>.
- [6]. Y. Azhar, A. E. Minarno, Y. Munarko, and Z. Ibrahim, "Image Retrieval Based on Texton Frequency-Inverse Image Frequency," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, pp. 87–94, May 2020, doi: 10.22219/kinetik.v5i2.1026.
- [7]. M. I. Thusnavis Bella and A. Vasuki, "An efficient image retrieval framework using fused information feature," *Computers and Electrical Engineering*, vol. 75, pp. 46–60, May 2019, doi: 10.1016/j.compeleceng.2019.01.022.
- [8]. S. Chavda and M. Goyani, "ROBUST CONTENT-BASED IMAGE RETRIEVAL USING ICCV, GLCM, AND DWT-MSLBP DESCRIPTORS," *Computer Science*, vol. 23, no. 1, pp. 5–36, 2022, doi: 10.7494/csci.2022.23.1.3821.
- [9]. M. B. Rajesh and S. Sathiamoorthy, "Co-occurrence of Edges and Valleys with Support Vector Machine for Content Based Image retrieval," in *Proceedings of the 4th International Conference on Computing Methodologies and Communication, ICCMC 2020*, Institute of Electrical and Electronics Engineers Inc., Mar. 2020, pp. 465–470. doi: 10.1109/ICCMC48092.2020.ICCMC-00087.
- [10]. A. Starczewski, P. Goetzen, and M. J. Er, "A New Method for Automatic Determining of the DBSCAN Parameters," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 10, no. 3, pp. 209–221, Jul. 2020, doi: 10.2478/jaiscr-2020-0014.
- [11]. Y. Wang, Y. Gu, and J. Shun, "Theoretically-Efficient and Practical Parallel DBSCAN," in *Proceedings of the ACM SIGMOD International Conference on Management of Data*, Association for Computing Machinery, Jun. 2020, pp. 2555–2571. doi: 10.1145/3318464.3380582.
- [12]. H. Liu, Y. Liu, Z. Qin, R. Zhang, Z. Zhang, and L. Mu, "A Novel DBSCAN Clustering Algorithm via Edge Computing-Based Deep Neural Network Model for Targeted Poverty Alleviation Big Data," *Wirel Commun Mob Comput*, vol. 2021, 2021, doi: 10.1155/2021/5536579.

- [13]. S. S. Li, "An Improved DBSCAN Algorithm Based on the Neighbor Similarity and Fast Nearest Neighbor Query," *IEEE Access*, vol. 8, pp. 47468–47476, 2020, doi: 10.1109/ACCESS.2020.2972034.
- [14]. Y. Chen, L. Zhou, N. Bouguila, C. Wang, Y. Chen, and J. Du, "BLOCK-DBSCAN: Fast clustering for large scale data," *Pattern Recognit*, vol. 109, Jan. 2021, doi: 10.1016/j.patcog.2020.107624.





UNIVERSITAS MUHAMMADIYAH MALANG



FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Muhammad Rivaldi Asyhari
NIM : 201910370311370
Judul TA : Enhancing Texture-Based Image Retrieval using GLCM and DBSCAN on a Multifaceted Dataset

Hasil Cek Plagiarisme dengan Turnitin

Table with 4 columns: No., Komponen Pengecekan, Nilai Maksimal Plagiarisme (%), Hasil Cek Plagiarisme (%)*. Rows include Bab 1-5 and Makalah Tugas Akhir.

*) Hasil cek plagiarism diisi oleh pemeriksa (staf TU)

*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui,

Pemeriksa (Staff TU)

Signature and official stamp of the examiner.



Kampus I
Jl. Bandung 1 Malang Jawa Timur
P +62 341 551 253 (Hunting)
F +62 341 460 435

Kampus II
Jl. Bendungan Sutarni No 188 Malang, Jawa Timur
P +62 341 551 149 (Hunting)
F +62 341 582 060

Kampus III
Jl. Raya Tlogomas No 246 Malang Jawa Timur
P +62 341 464 318 (Hunting)
F +62 341 460 435
E. webmaster@umm.ac.id