

PENERAPAN KLASIFIKASI ALGORITMA C4.5 PADA FITUR GRAY LEVEL CO-OCCURRANCE MATRIX UNTUK ANALISA TEKSTUR CITRA WAJAHIlham Kurniawan¹; Dwiza Riana²

^{1,2}Magister Ilmu Komputer
STMIK Nusa Mandiri Jakarta
www.nusamandiri.ac.id
ilhamkurniawan2100@gmail.com, dwiza@nusamandiri.ac.id



Abstract— Research on facial images is useful to distinguish the characteristics of each human being. The introduction of healthy and unhealthy facial skin images aims to identify human skin types automatically. For this purpose features such as contrast, correlation, energy, homogeneity, which are features of the Gray Level Co-occurrence Matrices (GLCM) are used. This study proposes a method for analyzing and classifying the GLCM texture on facial skin. The image used in this study was taken in the face image section which consists of the skin of the cheek and the whole face. The methods used are image acquisition, facial skin image, ROI selection, RGB image conversion to gray image, GLCM feature extraction, C4.5 algorithm classification and evaluation. The results showed that the C4.5 algorithm classification on texture analysis of facial images produced an accuracy value of 66.67%, the accuracy value was still low and the need for further research could not be used to increase the accuracy of texture analysis of facial images.

Keywords: Gray Level Co-Occurrence Matrix, Digital Image, Face Image, C4.5 Algorithm.

Intisari—Penelitian tentang citra wajah berguna untuk membedakan ciri-ciri dari setiap manusia. Pengenalan terhadap citra kulit wajah sehat dan tidak sehat bertujuan untuk mengidentifikasi jenis kulit manusia secara otomatis. Untuk tujuan tersebut fitur-fitur seperti contrast, correlation, energy, homogeneity, yang merupakan fitur Gray Level Co-occurrence Matrices (GLCM) digunakan. Penelitian ini mengusulkan metode untuk analisa dan klasifikasi tekstur GLCM pada kulit wajah. Citra yang digunakan pada penelitian ini diambil pada bagian citra wajah yang terdiri dari kulit pipi dan wajah keseluruhan. Metode yang digunakan yaitu akuisisi citra, citra kulit wajah, seleksi ROI, konversi citra RGB ke citra abu-abu, ekstraksi fitur GLCM, klasifikasi algoritma C4.5 dan evaluasi. Hasil penelitian menunjukkan bahwa klasifikasi algoritma C4.5 pada analisa tekstur citra wajah menghasilkan nilai akurasi sebesar 66,67%, nilai akurasi tersebut masih rendah dan belum dapat digunakan perlunya penelitian lebih lanjut untuk meningkatkan nilai akurasi analisa tekstur citra wajah.

Kata Kunci: Gray Level Co-occurrence matrix, citra digital, citra wajah, algoritma C4.5.

PENDAHULUAN

Orang semakin khawatir tentang kesehatan dan kecantikan mereka sendiri dalam beberapa tahun terakhir. Ada hubungan langsung antara keadaan kulit dan kecantikan kesehatan. Permukaan kondisi kulit tergantung pada karakteristik tekstur kulit. Setiap objek dengan bentuk fisik memiliki tekstur yang unik. Dengan kata lain, objek yang berbeda memiliki tekstur yang berbeda, sehingga tekstur adalah karakter penting bagi orang untuk mengamati dan mengidentifikasi objek (C. Yang, Liu, & Lan, 2018)(Pang, Chen, Wang, Chang, & Shao, 2017). Kulit adalah organ terbesar dari tubuh manusia

yang terdiri dari warna kulit yang berbeda. Analisis tekstur kulit memainkan peran penting dalam menilai kesehatan kulit dan untuk diagnosis gangguan kulit (C. Yang et al., 2018). Kondisi kulit wajah, sebagai sinyal kesehatan yang berubah dengan cepat, memberikan informasi yang berguna tentang status kesehatan dan fisiologis individu saat ini (Wei, Tiddeman, & Stephen, 2018).

Gray level co-occurrence matrices (GLCM), dikembangkan oleh Haralick et al. (1973), adalah metodologi pemrosesan gambar yang digunakan untuk menggambarkan hubungan spasial antara nilai-nilai abu-abu dalam gambar 2D. Karya asli Haralick et al. (1973) melangkah lebih jauh untuk menunjukkan tingkat penerapannya pada

photomicrographs level abu-abu dari serangkaian sampel batu pasir. Sejak itu, GLCM telah digunakan secara luas dalam berbagai aplikasi (Jardine, Miller, & Becker, 2018). Dari semua teknik analisis tekstur, saat ini mungkin yang paling banyak digunakan adalah yang didasarkan pada algoritma gray co-occurrence matrix (GLCM). Metode GLCM pertama kali diusulkan oleh Haralick et al. (1973) sebagai cara untuk mengklasifikasikan gambar menggunakan pengukuran statistik urutan kedua (Pantic, Dimitrijevic, Nestic, & Petrovic, 2016).

Algoritma C4.5 adalah salah satu algoritma metode *decision tree* yang paling efisien, kuat dan populer yang digunakan untuk tugas klasifikasi data medis, maka algoritma ini dipilih di antara semua algoritma lain dari metode *decision tree* (Lakshmi, Indumathi, & Ravi, 2016).

Penelitian mengenai pengolahan citra sudah banyak dilakukan sebelumnya oleh beberapa peneliti dan dalam berbagai macam bidang, seperti penelitian (Kurniawan & Riana, 2018), pengenalan pola (Liu, Wen, Gao, Gao, & Nie, 2018), ekstraksi fitur (P. Yang & Yang, 2016), pencitraan biomarker (Tran, Childs, Probst, Farhat, & Czarnota, 2018), klasifikasi kanker payudara (Mohammed et al., 2018), seleksi fitur, klasifikasi citra tunggal (Ramdhani & Riana, 2017), deteksi tepi (Riana, Dewi, Widyantoro, & Mengko, 2014), peningkatan kualitas citra (Surya Gowri, Amudha, Gowri, & Amudha, 2014), otsu threshold (Azeroual & Afdel, 2017), robust edge detection (Biswas & Hazra, 2018).

BAHAN DAN METODE

Data citra yang digunakan pada penelitian ini terdiri dari citra digital pipi dan citra digital keseluruhan wajah. Contoh citra-citra yang digunakan dalam penelitian ini dapat dilihat pada Gambar 1

Citra-citra pada penelitian ini diperoleh dari citra digital yang diambil dari klinik kecantikan yang dapat diakses secara *online* dan dapat digunakan secara bebas. Jumlah citra yang digunakan dapat dilihat pada Tabel 1

Dalam metode yang diusulkan seperti yang ditunjukkan pada Gambar 1, gambar yang diperoleh mengalami tahap pra-pemrosesan yang terdiri dari dua tahapan yaitu *contrast stretching* untuk peningkatan citra dan penyaringan *noise* oleh *filter low pass filter*. wilayah citra kemudian dipilih yang merupakan citra *Red Green Blue* (RGB) yang kemudian diubah menjadi citra abu-abu. *GLCM* dari gambar abu-abu kemudian ditemukan fitur seperti *covariance*, *standard deviation*, *energy*, dan *homogeneity* diekstraksi.

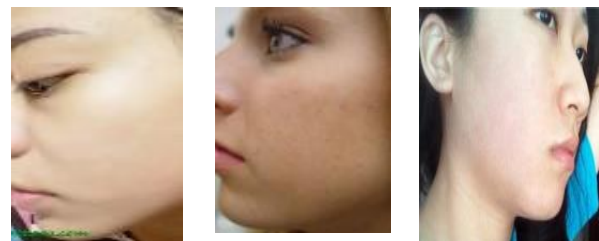
Berdasarkan gambar bagan tahapan penelitian di atas, dapat dijelaskan sebagai berikut:

1. Akuisisi citra digital adalah penciptaan representasi yang dikodekan secara digital dari karakteristik visual suatu objek, seperti pemandangan fisik atau struktur interior suatu objek (He & Zhang, 2018).

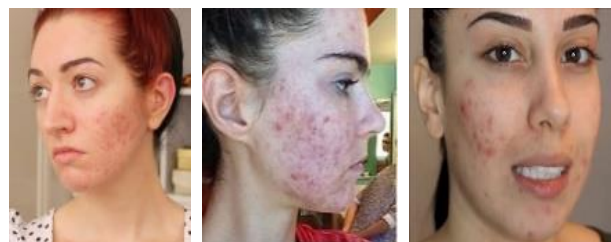


Sumber: (Kurniawan & Riana, 2019)
Gambar1. Bagan Tahapan Penelitian

2. Tekstur kulit wajah sehat dan tidak sehat dipilih untuk dianalisa.



Sumber: (Kurniawan & Riana, 2019)
Gambar 2. Citra Kulit Wajah Sehat



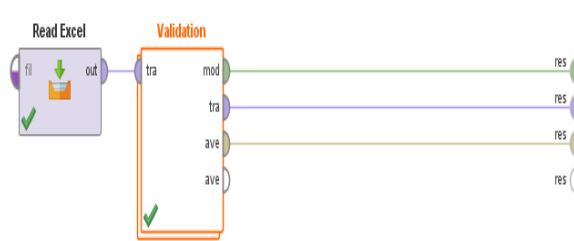
Sumber: (Kurniawan & Riana, 2019)
Gambar 3. Citra Kulit Wajah Tidak Sehat

3. *Region of Interest (ROI) selection*, adalah sampel dalam kumpulan data yang diidentifikasi untuk tujuan tertentu (Gaidel, 2017).

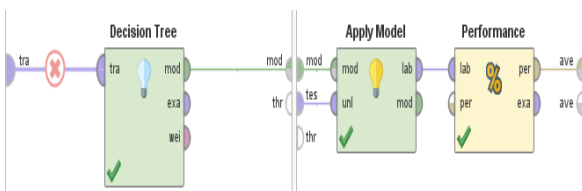


Sumber: (Kurniawan & Riana, 2019)
Gambar 4. Hasil Seleksi ROI

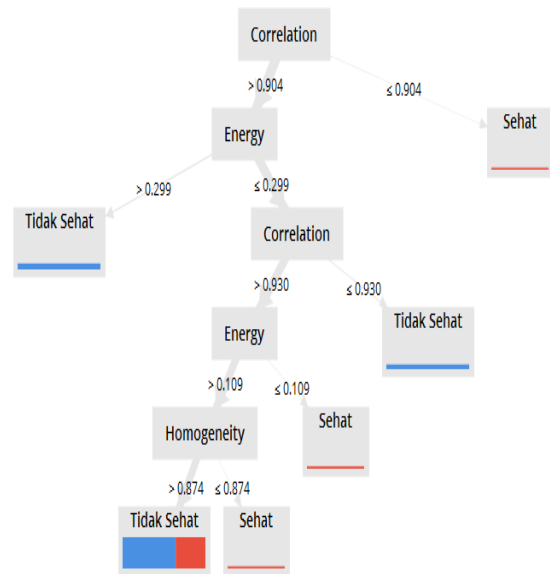
4. Citra dikonversi dari citra berwarna menjadi citra abu-abu untuk dianalisa dengan ekstraksi fitur GLCM.
5. Ekstraksi fitur adalah proses pengambilan ciri sebuah objek yang dapat menggambarkan karakteristik dari objek tersebut (Satria & Mushthofa, 2013).
6. Hasil ekstraksi fitur GLCM di klasifikasikan menggunakan klasifikasi algoritma pohon keputusan C4.5 dengan *machine learning*. Pada penelitian ini, pembuatan model menggunakan algoritma C4.5 yang akan disajikan pada gambar 5, metode validasi dalam *modelling* dilakukan dengan *cross validation* dimana, data akan menghasilkan sebuah algoritma, pohon keputusan, dan tingkat akurasi data. Hasil dari model tersebut akan di gambarkan pada gambar 6 dan 7.



Sumber: (Kurniawan & Riana, 2019)
Gambar 5. Proses pemodelan data



Sumber: (Kurniawan & Riana, 2019)
Gambar 6. Pemodelan data dengan cross validation



Sumber: (Kurniawan & Riana, 2019)
Gambar 7. Hasil pohon keputusan algoritma C4.5

7. Hasil di evaluasi untuk kepentingan penelitian lebih lanjut.

GLCM gambar abu-abu juga ditemukan menggunakan rumus:

$$Covariance : C(x, y) = E\{[x - E\{x\}] - [y - E\{y\}] \} \quad (1)$$

$$Standard Deviation : s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x - x)^2} \quad (2)$$

di mana n = jumlah total piksel di wilayah yang dipilih.

$$Homogeneity = \sum_{ij} \frac{p(i, j)}{1 + |i - j|} \quad (3)$$

di mana p (i, j) adalah entri (i, j) ketergantungan spasial keabuan

$$Matrix Energy : \sum_{ij} p(i, j)^2 \quad (4)$$

di mana p (i, j) adalah entri (i, j) dalam matriks ketergantungan spasial keabuan.

HASIL DAN PEMBAHASAN

Gambar 2 menunjukkan contoh sampel citra kulit wajah sehat dan Gambar 3 menunjukkan citra sampel kulit wajah tidak sehat, sedangkan Tabel 1 menunjukkan hasil ekstraksi nilai GLCM.

Tabel Jika anda menyajikan tabel harus menyebutkan nama table dan sebutkan sumber tabel, seperti contoh dibawah ini:

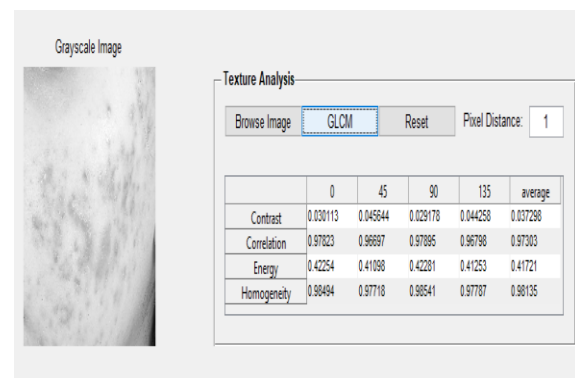
Tabel 1. Hasil Ekstraksi Fitur GLCM

Image s	Parameters	0	45	90	135	Average
Sampe l 1	Contrast	0.077	0.099	0.073	0.094	0.086
	Correlati on	0.986	0.982	0.987	0.983	0.984
	Energy	0.149	0.144	0.149	0.145	0.147
	Homoge neity	0.961	0.951	0.963	0.954	0.957
	Contrast	0.071	0.105	0.084	0.106	0.091
Sampe l 2	Contrast	0.515	0.02	0.583	0.33	0.862
	Correlati on	0.982	0.974	0.979	0.974	0.977
	Energy	0.302	0.297	0.299	0.295	0.298
	Homoge neity	0.968	0.956	0.961	0.954	0.960
	Contrast	0.049	0.068	0.041	0.064	0.056
Sampe l 3	Contrast	0.892	0.724	0.453	0.739	0.202
	Correlati on	0.981	0.975	0.984	0.976	0.979
	Energy	0.242	0.234	0.242	0.236	0.239
	Homoge neity	0.978	0.969	0.979	0.971	0.974
	Contrast	0.034	0.051	0.034	0.046	0.041
Sampe l 4	Contrast	0.567	0.443	0.638	0.366	0.753
	Correlati on	0.987	0.981	0.987	0.983	0.985
	Energy	0.229	0.222	0.228	0.224	0.226
	Homoge neity	0.982	0.974	0.982	0.976	0.979
	Contrast	0.127	0.148	0.122	0.182	0.145
Sampe l 5	Contrast	0.86	0.87	0.10	0.49	0.33
	Correlati on	0.961	0.954	0.963	0.944	0.955
	Energy	0.201	0.193	0.202	0.190	0.196
	Homoge neity	0.944	0.934	0.947	0.929	0.939
	Contrast	0.134	0.182	0.123	0.187	0.156
Sampe l 6	Contrast	0.30	0.39	0.35	0.22	0.82
	Correlati on	0.975	0.966	0.977	0.965	0.971
	Energy	0.259	0.250	0.258	0.250	0.254
	Homoge neity	0.954	0.940	0.955	0.941	0.948
	Contrast	0.158	0.221	0.136	0.210	0.181
Sampe l 7	Contrast	0.57	0.59	0.08	0.20	0.61
	Correlati on	0.959	0.943	0.965	0.946	0.954
	Energy	0.253	0.241	0.254	0.242	0.248
	Homoge neity	0.938	0.922	0.941	0.923	0.931
	Contrast	0.163	0.222	0.147	0.223	0.189
Sampe l 8	Contrast	0.31	0.27	0.94	0.91	0.36

Image s	Parameters	0	45	90	135	Average
Sampe l 1	Correlati on	0.969	0.958	0.972	0.958	0.965
	Energy	0.283	0.272	0.283	0.272	0.277
	Homoge neity	0.935	0.920	0.937	0.916	0.927
	Contrast	0.078	0.115	0.081	0.106	0.095
	Correlati on	0.992	0.989	0.992	0.989	0.990
Sampe l 9	Energy	0.185	0.181	0.185	0.182	0.183
	Homoge neity	0.967	0.958	0.966	0.959	0.962
	Contrast	0.076	0.109	0.078	0.103	0.091
	Correlati on	0.993	0.990	0.993	0.990	0.991
	Energy	0.192	0.189	0.192	0.189	0.191
Sampe l 10	Homoge neity	0.968	0.960	0.968	0.961	0.964
	Contrast	0.063	0.37	0.053	0.34	0.705
	Correlati on	0.993	0.990	0.993	0.990	0.991
	Energy	0.192	0.189	0.192	0.189	0.191
	Homoge neity	0.968	0.960	0.968	0.961	0.964

Sumber: (Kurniawan & Riana, 2019)

Hasil yang diperoleh menunjukkan bahwa fitur ekstraksi GLCM tidak dapat memahami pola tekstur kulit. Hasil untuk perhitungan fitur GLCM akan disajikan digambar 8.



Sumber: (Kurniawan & Riana, 2019)

Gambar 8. Hasil perhitungan GLCM

KESIMPULAN

Hasil klasifikasi algoritma C4.5 pada fitur GLCM untuk analisa tekstur citra wajah menghasilkan nilai akurasi pada software RapidMiner Studio sebesar 66,67 %, nilai akurasi tersebut masih rendah dan belum dapat digunakan untuk klasifikasi tekstur citra wajah, perlunya ditambahkan data menjadi lebih banyak lagi, ataupun dengan mengganti metode seperti metode Naive Bayes Classifier, ataupun dilakukan optimasi untuk meningkatkan nilai akurasi seperti Particle Swarm Optimization.

REFERENSI

- Azeroual, A., & Afdel, K. (2017). Fast Image Edge Detection based on Faber Schauder Wavelet and Otsu Threshold. *Heliyon*, 3(12), 1–19. <https://doi.org/10.1016/j.heliyon.2017.e00485>
- Biswas, S., & Hazra, R. (2018). Robust edge detection based on Modified Moore-Neighbor. *Optik*, (2010). <https://doi.org/10.1016/j.ijleo.2018.05.011>
- Gaidel, A. (2017). Method of automatic ROI selection on lung CT images. *Procedia Engineering*, 201, 258–264. <https://doi.org/10.1016/j.proeng.2017.09.612>
- He, X., & Zhang, J. (2018). Emerging market MNCs' cross-border acquisition completion: Institutional image and strategies. *Journal of Business Research*, (April), 0–1. <https://doi.org/10.1016/j.jbusres.2018.04.014>
- Jardine, M. A., Miller, J. A., & Becker, M. (2018). Coupled X-ray computed tomography and grey level co-occurrence matrices as a method for quantification of mineralogy and texture in 3D. *Computers and Geosciences*, 111, 105–117. <https://doi.org/10.1016/j.cageo.2017.11.005>
- Kurniawan, I., & Riana, D. (2018). Analisa Tekstur Kulit Wajah Menggunakan Fitur Gray Level Co-Occurance Matrix. *Seminar Nasional Inovasi Dan Tren (SNIT) 2018*, 187–192.
- Kurniawan, I., & Riana, D. (2019). *Laporan Akhir Penelitian Mandiri*. Jakarta.
- Lakshmi, B. N., Indumathi, T. S., & Ravi, N. (2016). A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy. *Procedia Technology*, 24, 1542–1549. <https://doi.org/10.1016/j.protcy.2016.05.128>
- Liu, Y., Wen, K., Gao, Q., Gao, X., & Nie, F. (2018). SVM based multi-label learning with missing labels for image annotation. In *Pattern Recognition* (Vol. 78). Elsevier Ltd. <https://doi.org/10.1016/j.patcog.2018.01.022>
- Mohammed, M. A., Al-khateeb, B., Rashid, A. N., Ibrahim, D. A., Ghani, M. K. A., & Mostafa, S. A. (2018). Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images. *Computers and Electrical Engineering*, 70(August), 871–882. <https://doi.org/10.1016/j.compeleceng.2018.01.033>
- Pang, H., Chen, T., Wang, X., Chang, Z., & Shao, S. (2017). Quantitative evaluation methods of skin condition based on texture feature parameters. *Saudi Journal of Biological Sciences*. <https://doi.org/10.1016/j.sjbs.2017.01.021>
- Pantic, I., Dimitrijevic, D., Nestic, D., & Petrovic, D. (2016). Gray level co-occurrence matrix algorithm as pattern recognition biosensor for oxidopamine-induced changes in lymphocyte chromatin architecture. *Journal of Theoretical Biology*, 406, 124–128. <https://doi.org/10.1016/j.jtbi.2016.07.018>
- Ramdhani, Y., & Riana, D. (2017). Hierarchical Decision Approach based on Neural Network and Genetic Algorithm method for single image classification of Pap smear. *2017 Second International Conference on Informatics and Computing (ICIC)*, 1–6. <https://doi.org/10.1109/IAC.2017.8280587>
- Riana, D., Dewi, D. E. O., Widyantoro, D. H., & Mengko, T. L. R. (2014). Color canals modification with canny edge detection and morphological reconstruction for cell nucleus segmentation and area measurement in normal Pap smear images. *AIP Conference Proceedings*, 1589, 414–417. <https://doi.org/10.1063/1.4868832>
- Satria, D., & Mushthofa. (2013). Perbandingan Metode Ekstraksi Ciri Histogram dan PCA untuk Mendeteksi Stoma pada Citra Penampang Daun Freycinetia. *Jurnal Ilmu Komputer Agri-Informatika*, 2, 20–28.
- Surya Gowri, D., Amudha, T., Gowri, D. S., & Amudha, T. (2014). A Review on Mammogram Image Enhancement Techniques for Breast Cancer Detection. *2014 International Conference on Intelligent Computing Applications*, 47–51. <https://doi.org/10.1109/ICICA.2014.19>
- Tran, W. T., Childs, C., Probst, H., Farhat, G., & Czarnota, G. J. (2018). Imaging Biomarkers for Precision Medicine in Locally Advanced

- Breast Cancer. *Journal of Medical Imaging and Radiation Sciences*, 1–10.
<https://doi.org/10.1016/j.jmir.2017.09.006>
- Wei, K., Tiddeman, B., & Stephen, I. D. (2018). Evolution and Human Behavior Skin texture and colour predict perceived health in Asian faces. *Evolution and Human Behavior*, 39(3), 320–335.
<https://doi.org/10.1016/j.evolhumbehav.2018.02.003>
- Yang, C., Liu, H., & Lan, Z. (2018). Optik Simultaneous texture image enhancement and directional field estimation based on local quality metrics. *Optik - International Journal for Light and Electron Optics*, 158, 1203–1219.
<https://doi.org/10.1016/j.ijleo.2017.12.054>
- Yang, P., & Yang, G. (2016). Feature extraction using dual-tree complex wavelet transform and gray level co-occurrence matrix. *Neurocomputing*, 197, 212–220.
<https://doi.org/10.1016/j.neucom.2016.02.061>