

PENERAPAN KLASIFIKASI ALGORITMA C4.5 PADA FITUR GRAY LEVEL CO-OCCURRENCE MATRIX UNTUK ANALISA TEKSTUR CITRA WAJAH**Ilham 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 karakterisasi 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, and Lan 2018)(Pang et al. 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, Liu, and Lan 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, and 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, and 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 et al. 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, and Ravi 2016).

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

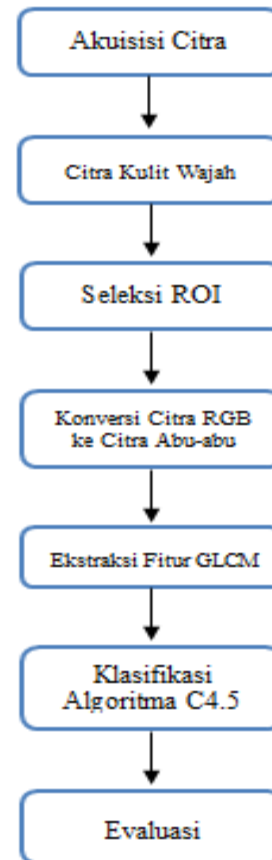
Citra-citra pada penelitian ini diperoleh dari citra digital yang diambil dari klinik kecantikan yang dapat diakses secara *online* dan dapat digunakan secara bebas. Jumlah citra yang digunakan dapat dilihat pada Tabel 1

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

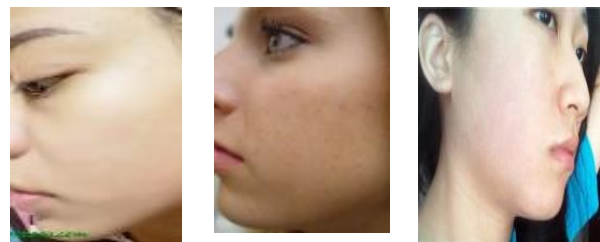
karakteristik visual suatu objek, seperti pemandangan fisik atau struktur interior suatu objek (He and Zhang 2018).



Sumber: (Kurniawan and Riana 2019)

Gambar1. Bagan Tahapan Penelitian

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



Sumber: (Kurniawan and Riana 2019)

Gambar 2. Citra Kulit Wajah Sehat



Sumber: (Kurniawan and Riana 2019)

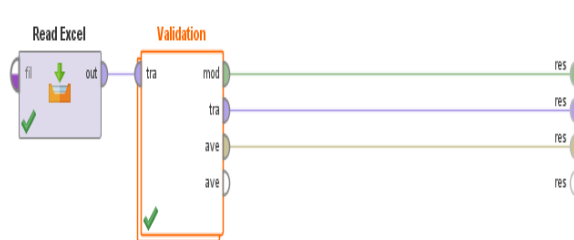
Gambar 3. Citra Kulit Wajah Tidak Sehat

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

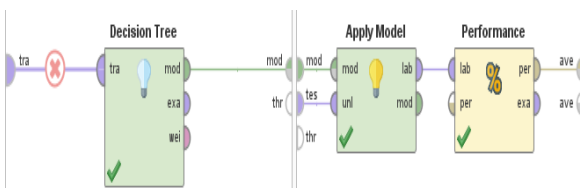


Sumber: (Kurniawan and Riana 2019)
Gambar 4. Hasil Seleksi ROI

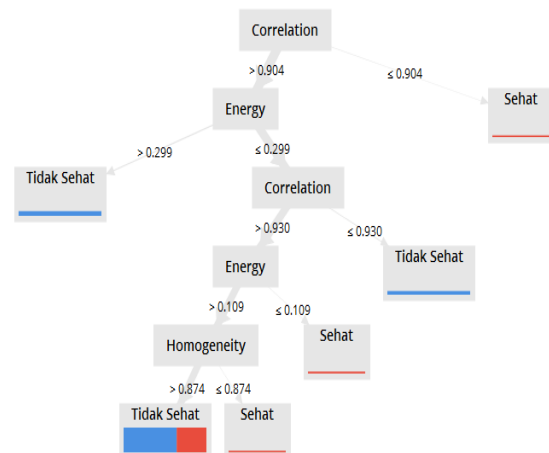
4. Citra dikonversi dari citra berwarna menjadi citra abu-abu untuk dianalisa dengan ekstraksi fitur GLCM.
5. Ekstraksi fitur adalah proses pengambilan ciri sebuah objek yang dapat menggambarkan karakteristik dari objek tersebut (Satria and Mushthofa 2013).
6. Hasil ekstraksi fitur GLCM di klasifikasikan 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 and Riana 2019)
Gambar 5. Proses pemodelan data



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



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

Jelaskan gambar 7 diatas

7. Hasil di evaluasi untuk kepentingan penelitian lebih lanjut.

GLCM gambar abu-abu juga ditemukan menggunakan rumus:

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

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

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

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

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

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

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

Images	Parameters	0	45	90	135	Average
Sample 1		0.077386	0.099491	0.073702	0.094504	0.086271
	Correlation	0.98645	0.98259	0.98709	0.98338	0.98488
	Energy	0.14913	0.14410	0.14996	0.14541	0.14715
	Homogeneity	0.96185	0.95142	0.96343	0.95401	0.95768
	Contrast	0.071515	0.10502	0.084583	0.10633	0.091862
Sample 2		0.07165	0.10534	0.08440	0.10603	0.09160
	Correlation	0.98265	0.97434	0.97940	0.97403	0.97760
	Energy	0.30293	0.29747	0.29990	0.29558	0.29897
	Homogeneity	0.96804	0.95691	0.96177	0.95430	0.96025
	Contrast	0.049892	0.068724	0.041453	0.064739	0.056202
Sample 3		0.04989	0.06802	0.04196	0.06443	0.05657
	Correlation	0.98189	0.97502	0.98496	0.97643	0.97957
	Energy	0.24259	0.23463	0.24287	0.23628	0.23909
	Homogeneity	0.97891	0.96950	0.97927	0.97149	0.97479
	Contrast	0.034567	0.051443	0.034638	0.046366	0.041753
Sample 4		0.03463	0.05155	0.03460	0.04639	0.04104
	Correlation	0.98763	0.98155	0.98760	0.98339	0.98504
	Energy	0.22910	0.22253	0.22887	0.22456	0.22627
	Homogeneity	0.98272	0.97429	0.98268	0.97684	0.97913
	Contrast	0.12786	0.14887	0.12210	0.18249	0.14533
Sample 5		0.12716	0.14878	0.12201	0.18258	0.14588
	Correlation	0.96116	0.95478	0.96301	0.94458	0.95588
	Energy	0.20109	0.19306	0.20259	0.19056	0.19682
	Homogeneity	0.94473	0.93457	0.94751	0.92928	0.93902
	Contrast	0.13430	0.18239	0.12335	0.18722	0.15682
Sample 6		0.13454	0.18274	0.12365	0.18782	0.15644
	Correlation	0.97554	0.96674	0.97765	0.96582	0.97144
	Energy	0.25941	0.25051	0.25877	0.25064	0.25483
	Homogeneity	0.95433	0.94080	0.95573	0.94125	0.94803
	Contrast	0.15857	0.22159	0.13608	0.21020	0.18161
Sample 7		0.15897	0.22186	0.13664	0.21077	0.18106
	Correlation	0.95997	0.94386	0.96564	0.94677	0.95406
	Energy	0.25323	0.24162	0.25474	0.24292	0.24813
	Homogeneity	0.93897	0.92238	0.94190	0.92386	0.93178
	Contrast	0.16331	0.22227	0.14794	0.22391	0.18936

untuk meningkatkan nilai akurasi seperti Particle Swarm Optimazation.

Feature Parameters." *Saudi Journal of Biological Sciences*.

REFERENSI

- Azeroual, Assma, and Karim Afdel. 2017. "Fast Image Edge Detection Based on Faber Schauder Wavelet and Otsu Threshold." *Heliyon* 3(12): 1–19.
- Biswas, Soumen, and Ranjay Hazra. 2018. "Robust Edge Detection Based on Modified Moore-Neighbor." *Optik* (2010).
- Gaidel, Andrey. 2017. "Method of Automatic ROI Selection on Lung CT Images." *Procedia Engineering* 201: 258–64.
- He, Xinming, and Jianhong Zhang. 2018. "Emerging Market MNCs' Cross-Border Acquisition Completion: Institutional Image and Strategies." *Journal of Business Research* (April): 0–1.
- Jardine, M. A., J. A. Miller, and M. Becker. 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: 105–17.
- Kurniawan, Ilham, and Dwiza Riana. 2018. "Analisa Tekstur Kulit Wajah Menggunakan Fitur Gray Level Co-Occurrence Matrix." *Seminar Nasional Inovasi dan Tren (SNIT) 2018*: 187–92.
- . 2019. *Laporan Akhir Penelitian Mandiri*. Jakarta.
- Lakshmi, B.N., T.S. Indumathi, and Nandini Ravi. 2016. "A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy." *Procedia Technology* 24: 1542–49.
- Liu, Yang et al. 2018. *78 Pattern Recognition SVM Based Multi-Label Learning with Missing Labels for Image Annotation*. Elsevier Ltd.
- Mohammed, Mazin Abed et al. 2018. "Neural Network and Multi-Fractal Dimension Features for Breast Cancer Classification from Ultrasound Images." *Computers and Electrical Engineering* 0: 1–12.
- Pang, Hui et al. 2017. "Quantitative Evaluation Methods of Skin Condition Based on Texture
- Pantic, Igor, Draga Dimitrijevic, Dejan Nesic, and Danica Petrovic. 2016. "Gray Level Co-Occurrence Matrix Algorithm as Pattern Recognition Biosensor for Oxidopamine-Induced Changes in Lymphocyte Chromatin Architecture." *Journal of Theoretical Biology* 406: 124–28.
- Ramdhani, Yudi, and Dwiza Riana. 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.
- Riana, Dwiza, Dyah Ekashanti Octorina Dewi, Dwi H. Widyantoro, and Tati Latifah R Mengko. 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: 414–17.
- Satria, Dony, and 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, T. Amudha, D. Surya Gowri, and T. Amudha. 2014. "A Review on Mammogram Image Enhancement Techniques for Breast Cancer Detection." *2014 International Conference on Intelligent Computing Applications*: 47–51.
- Tran, William T. et al. 2018. "Imaging Biomarkers for Precision Medicine in Locally Advanced Breast Cancer." *Journal of Medical Imaging and Radiation Sciences*: 1–10.
- Wei, Kok, Bernard Tiddeman, and Ian D Stephen. 2018. "Evolution and Human Behavior Skin Texture and Colour Predict Perceived Health in Asian Faces." *Evolution and Human Behavior* 39(3): 320–35.
- Yang, Chao, Hong Liu, and Zengmei Lan. 2018. "Optik Simultaneous Texture Image Enhancement and Directional Field Estimation Based on Local Quality Metrics." *Optik - International Journal for Light and Electron Optics* 158: 1203–19.

Yang, Peng, and Guowei Yang. 2016. "Feature Extraction Using Dual-Tree Complex Wavelet Transform and Gray Level Co-Occurrence Matrix." *Neurocomputing* 197: 212–20.