

IDENTIFIKASI CITRA MASSA KISTIK BERDASAR FITUR GRAY- LEVEL CO-OCCURRENCE MATRIX

Hari Wibawanto¹⁾, Adhi Susanto²⁾, Thomas Sri Widodo³⁾, S. Maesadji Tjokronegoro⁴⁾

¹⁾ Fakultas Teknik Universitas Negeri Semarang
Gd. E6 Kampus Sekaran Gunungpati Semarang

^{2,3)} Fakultas Teknik Universitas Gadjahmada Jogjakarta
Jl. Grafika Yogyakarta

⁴⁾ Fakultas Kedokteran Universitas Gadjahmada Yogyakarta
Jln. Farmako, Sekip Utara, Jogjakarta, 55281
E-mail: hariwibawanto@gmail.com

ABSTRACT

We have studied the effectiveness of using texture features derived from gray-level co-occurrence matrix (GLCM) matrices for classification of cystic mass and non-cystic mass in ultra sonograms. Twenty-three (23) region of interest (ROIs) containing cystic masses and fifty-five (55) non-cystic masses were extracted from ultra sonogram for this study. For each ROI of 50x50 pixels, seven features (energy, inertia, entropy, homogeneity, maximum probability, inverse difference moment, and correlation) were calculated. The importance of each feature in distinguishing cystic masses from non-cystic masses was determined by linear discriminant analysis with SPSS version 11.5 program. As a result of a study, it was found that all seven features can distinguishing cystic masses from non-cystic masses with an accuracy about 91 %-92.3%. Those levels of accuracy also found when two features (energy and inverse difference moment) was excluded from analysis. The result demonstrate the feasibility of using texture features based on GLCM for distinguishing cystic masses from non-cystic masses of ultra sonogram .

Keywords: Gray-level Co-occurrence Matrix Ultrasonografi, massa kistik, fitur tekstur, analisis tekstur, analisis diskriminan

1. PENDAHULUAN

Ultrasonografi banyak digunakan pada tahap diagnosis untuk mendapatkan citra organ bagian dalam tubuh manusia dengan cara memindaikan berkas isyarat ultrasonik pada objek. Hasil ultrasonografi berupa citra keabuan (*grayscale*) yang merupakan *echo* dari getaran ultrasonik yang dipantulkan oleh lapisan kulit dan organ bagian dalam. Citra semacam itu biasanya diinterpretasi oleh ahli yang belajar secara khusus, umumnya tenaga medis. Kekeliruan dalam memahami citra ultrasonogram bisa berakibat fatal, karena hasil interpretasi citra tersebut menjadi dasar bagi tindakan medis selanjutnya.

Hakikatnya, ada perbedaan antara citra normal dengan dan citra abnormal karena adanya 'sesuatu' di dalam objek yang dipindai, dalam hal ini tubuh manusia. Pada pemindaian dengan ultrasonograf, citra massa kistik memberikan penampakan berbeda dengan massa solid. Massa kistik secara sederhana dapat dikatakan sebagai kantung berisi gas atau cairan (darah, nanah atau cairan lain). Kista atau kistoma, adalah penamaan bagi kantung berisi cairan yang umum terdapat di berbagai bagian tubuh manusia. Dengan beberapa pengecualian, kista atau kistoma yang berada di dalam dan disekitar organ reproduksi wanita, umumnya merupakan kelainan yang tidak berbahaya. Berbeda halnya dengan massa solid yang, juga dengan beberapa kekecualian, bisa jadi

merupakan perwujudan dari tumor yang cenderung ganas.

Cita massa kistik menampakkan wilayah yang relatif lebih hitam dibandingkan dengan wilayah sekelilingnya. Wilayah kehitaman ini juga tampak lebih merata atau homogen. Ini berbeda dengan citra massa solid yang umumnya berupa citra kehitaman tetapi kurang atau tidak merata (Palmer, 1996: 32-35).

Teknik analisis citra berbantuan komputer berpotensi dapat membedakan adanya massa kistik atau massa solid tubuh manusia atas dasar kemampuannya membedakan intensitas dan homogenitas kehitaman dari citra hasil pemindaian ultrasonograf. Kemampuan komputer ini dapat dimanfaatkan untuk membantu proses analisis citra ultrasonogram dalam menentukan apakah citra hasil pemindaian merupakan citra massa kistik atau solid. Interpretasi citra oleh komputer dapat digunakan oleh dokter, setidaknya sebagai bahan diagnosis awal, atau bila akurasinya memadai dapat digunakan sebagai alat bantu utama dokter dalam menentukan jenis massa yang terdeteksi oleh ultrasonograf.

2. ANALISIS TEKSTUR

Tekstur adalah sifat atau properti yang dimiliki oleh permukaan benda. Pada lingkungan alami, permukaan objek seperti potongan kayu, rerumputan, hamparan pasir pantai, tekstil, kulit, dan sebagainya memiliki tekstur. Citra juga dapat

dipandang memiliki tekstur yang terbentuk akibat variasi dan atau gradasi keabuan pada citra digital. Tekstur adalah fitur yang bergantung konteks, maksudnya, tekstur tidak dapat didefinisikan hanya dari piksel saja tetapi harus dalam kaitannya dengan piksel lain dalam suatu wilayah citra (Hauta-Kasari, 1999: 13).

Metode analisis tekstur citra keabuan telah banyak diteliti. Tekstur sebenarnya sulit didefinisikan meskipun secara intuitif manusia mampu membedakan dua atau lebih citra yang berbeda penampakannya. Berbagai metode analisis tekstur diajukan, umumnya berdasar pada pendekatan statistis. Fitur-fitur yang digunakan dalam analisis tekstur secara statistis umumnya didasarkan pada *gray level cooccurrence matrices*, *gray level difference matrices*, *run length matrices*, spektra daya Fourier, fungsi autokorelasi, dan model *random-field*.

Salah satu metode analisis tekstur yang paling banyak digunakan adalah *gray level co-occurrence matrix* yang didasarkan pada statistika orde kedua. Matriks *cooccurrence* ini diperkenalkan pertama kali oleh Haralick (1973: 610) untuk mengekstrak fitur-fitur yang digunakan sebagai analisis citra hasil penginderaan jauh. *Co-occurrence* didefinisikan sebagai distribusi gabungan dari level keabuan dua piksel yang terpisah jarak dan arah tertentu ($\Delta x, \Delta y$). *Gray Level Cooccurrence Matrix (GLCM)*, atau Haralick et. al. (1973: 612) menyebutnya sebagai *Gray-Tone Spatial-Dependence Matrices*, P didefinisikan sebagai:

$$P(a,b) = \{ \# \text{ pair}(a,b) \mid \text{citra}(x,y) = a \Delta \text{citra}(x + \Delta x, y + \Delta y) = b \}$$

Dengan a, b adalah level keabuan.

3. GRAY-LEVEL CO-OCCURRENCE MATRIX

Pada analisis tekstur secara statistis, fitur tekstur dihitung berdasarkan distribusi statistik dari kombinasi intensitas piksel pada posisi tertentu relatif terhadap lainnya dalam suatu matriks citra. Bergantung pada jumlah piksel atau titik intensitas dalam masing-masing kombinasi, dibedakan adanya statistik orde-pertama, statistik orde-kedua dan statistik orde-lebih-tinggi (*higher-order statistics*). Metode GLCM (*gray-level co-occurrence matrix*) adalah salah satu cara mengekstrak fitur tekstur statistik orde-kedua (Albregtsen, 1995:1). GLCM (yang disebut juga *Grey Tone Spatial Dependency Matrix*) adalah tabulasi mengenai frekuensi atau seberapa seringnya kombinasi nilai kecerahan piksel yang berbeda posisinya terjadi dalam suatu citra (Hall-Beyer, 2007).

GLCM dihitung dengan algoritma sebagai berikut:

begin

forall gray levels $a, b \in \text{image}$

$P(a,b)=0;$

endfor

forall location $(x,y) \in \text{image}$ and $(x+\Delta x, y+\Delta y) \in \text{image}$

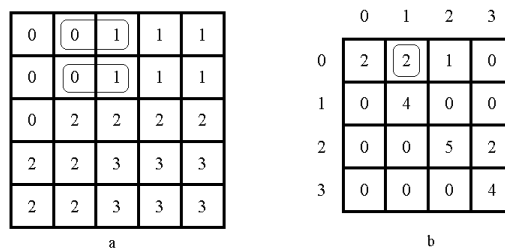
$a=\text{image}(x,y);$

$b=\text{image}(x+\Delta x, y+\Delta y);$

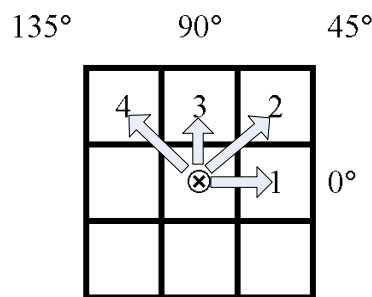
endfor

end

Berikut ini adalah gambaran pembentukan GLCM atas citra dengan 4 tingkat keabuan (gray level) pada jarak $d=1$ dan arah 0° .



Gambar 1. a. Contoh citra dengan 4 tingkat keabuan
b. GLCM pada jarak 1 arah 0°



Gambar 2. Dari piksel di tengah (\otimes) piksel 1 menunjukkan arah = 0° dengan jarak $d=1$, piksel 2 arah = 45° dengan jarak $d=1$, piksel 3 arah = 90° dengan jarak $d=1$, dan piksel 4 arah = 135° dengan jarak $d=1$

4. FITUR GLCM

Haralick et. al (1973: 613) mengekstrak 14 fitur dari matriks *cooccurrence* tersebut, meskipun dalam banyak aplikasi hanya beberapa fitur yang banyak digunakan, antara lain:

$$\text{Energi} = \sum_{a,b} P_{\phi,d}^2(a,b)$$

$$\text{Entropi} = \sum_{a,b} P_{\phi,d}(a,b) \log_2 P_{\phi,d}(a,b)$$

$$\text{Max Probablility} = \max_{a,b} P_{\phi,d}(a,b)$$

$$\text{Inverse Diff. Moment} = \sum_{a,b;a \neq b} \frac{P_{\phi,d}(a,b)}{|a-b|^2}$$

$$\text{Kontras} = \sum_{a,b} |a-b|^2 P_{\phi,d}(a,b)$$

$$\text{Homogenitas} = \sum_a \sum_b \frac{1}{1+(a-b)^2} P_{\phi,d}(a,b)$$

$$\text{Inertia} = \sum_a \sum_b (a-b)^2 P_{\phi,d}(a,b)$$

$$\text{Korelasi} = \frac{\sum_{a,b} [(ab)P_{\phi,d}(a,b)] - \mu_x \mu_y}{\sigma_x \sigma_y}$$

dengan

$$\mu_x = \sum_a a \sum_b P_{\phi,d}(a,b)$$

$$\mu_y = \sum_b b \sum_a P_{\phi,d}(a,b)$$

$$\sigma_x = \sum_a (a - \mu_x)^2 \sum_b P_{\phi,d}(a,b)$$

$$\sigma_y = \sum_b (b - \mu_y)^2 \sum_a P_{\phi,d}(a,b)$$

Matriks co-occurrence menangkap sifat tekstur tetapi tidak secara langsung dapat digunakan sebagai alat analisis, misalnya membandingkan dua tekstur. Data ini harus disarikan lagi agar didapatkan angka-angka yang bisa digunakan untuk mengklasifikasi tekstur. Haralick (1973) mengusulkan 14 ukuran (atau ciri/fitur), tetapi Connors dan Harlow pada tahun 1980-an mengkaji bahwa dari 14 fitur yang diusulkan Haralick tersebut, hanya 5 diantaranya yang biasanya digunakan. Kelima fitur itu adalah: energi, entropi, korelasi, homogenitas, dan inersia (Kulak, 2002:62).

5. CARA PENELITIAN

6. HASIL DAN PEMBAHASAN

Dilakukan analisis terhadap 78 citra USG (23 citra massa kistik, 55 citra massa non-kistik) ukuran 50x50 piksel. Citra massa non kistik meliputi citra mioma (27 citra), dan citra tumor padat (28 citra). Ekstraksi fitur dilakukan dengan GLCM pada 4 arah (0°, 45°, 90°, 135°) dan jarak 1 piksel. Fitur yang diekstrak adalah energi, inersia, entropi, homogenitas, probabilitas maksimum, inverse

difference moment, dan korelasi. Klasifikasi dilakukan dengan analisis diskriminan menggunakan paket program SPSS versi 11.5.0. Hasilnya adalah sebagai berikut:

Tabel 1. Hasil Klasifikasi (jarak = 1, arah = 0°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results^{a,c}

	Kelas	Predicted Group Membership		Total
		0	1	
Original	Count 0	17	6	23
	1	1	54	55
	% 0	73.9	26.1	100.0
	1	1.8	98.2	100.0
Cross-validated ^a	Count 0	17	6	23
	1	1	54	55
	% 0	73.9	26.1	100.0
	1	1.8	98.2	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 91.0% of original grouped cases correctly classified.

c. 91.0% of cross-validated grouped cases correctly classified.

Tabel 2. Hasil Klasifikasi (jarak = 1, arah = 45°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results^{a,c}

	Kelas	Predicted Group Membership		Total
		0	1	
Original	Count 0	17	6	23
	1	0	55	55
	% 0	73.9	26.1	100.0
	1	.0	100.0	100.0
Cross-validated ^a	Count 0	17	6	23
	1	2	53	55
	% 0	73.9	26.1	100.0
	1	3.6	96.4	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 92.3% of original grouped cases correctly classified.

c. 89.7% of cross-validated grouped cases correctly classified.

Tabel 3. Hasil Klasifikasi (jarak = 1, arah = 90°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results^{a,c}

	Kelas	Predicted Group Membership		Total
		0	1	
Original	Count 0	17	6	23
	1	0	55	55
	% 0	73.9	26.1	100.0
	1	.0	100.0	100.0
Cross-validated ^a	Count 0	17	6	23
	1	2	53	55
	% 0	73.9	26.1	100.0
	1	3.6	96.4	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

b. 92.3% of original grouped cases correctly classified.

c. 89.7% of cross-validated grouped cases correctly classified.

Tabel 4. Hasil Klasifikasi (jarak = 1, arah = 135°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results^{b,c}

		Predicted Group Membership		Total
		0	1	
Original	Count	0	6	23
		1	55	55
	%	0	26.1	100.0
		1	100.0	100.0
Cross-validated ^a	Count	0	6	23
		1	54	55
	%	0	26.1	100.0
		1	98.2	100.0

- a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
- b. 92.3% of original grouped cases correctly classified.
- c. 91.0% of cross-validated grouped cases correctly classified.

Berdasarkan analisis statistik terlihat bahwa untuk jarak piksel = 1 pada keempat arah didapatkan bahwa antara 91,0% - 92,3% massa kistik dan massa non-kistik dapat dibedakan berdasarkan fitur-fitur GLCM energi, inersia, entropi, homogenitas, probabilitas maksimum, inverse, dan korelasi secara bersama-sama.

Koefisien fungsi diskriminan yang diperoleh dari hasil analisis (untuk keempat arah, pada jarak j = 1) ditunjukkan berturut pada tabel 5-8 berikut ini.

Tabel 5. Koefisien fungsi diskriminan (jarak = 1, arah = 0°)

Standardized Canonical Discriminant Function Coefficients

	Function
	1
Energi	-.321
Inertia	-.293
Entropi	2.073
Homogenitas	-1.033
Max Probability	-.107
Inverse	-.833
Korelasi	.829

Tabel 6. Koefisien fungsi diskriminan (jarak = 1, arah = 45°)

Standardized Canonical Discriminant Function Coefficients

	Function
	1
Energi	.048
Inertia	-.403
Entropi	2.412
Homogenitas	-1.503
Maxprob	-.110
Inverse	-.830
Correlation	.946

Tabel 7. Koefisien fungsi diskriminan (jarak = 1, arah = 90°)

Standardized Canonical Discriminant Function Coefficients

	Function
	1
Energi	.048
Inertia	-.403
Entropi	2.412
Homogenitas	-1.503
Maxprob	-.110
Inverse	-.830
Correlation	.946

Tabel 8. Koefisien fungsi diskriminan (jarak = 1, arah = 135°)

Standardized Canonical Discriminant Function Coefficients

	Function
	1
Energi	-.274
Inertia	.856
Entropi	-2.117
Homogenitas	1.804
Max Probability	.127
Inverse	.917
Korelasi	-.870

Berdasarkan data yang diperoleh, pada semua arah diperoleh fakta bahwa fitur entropi dan homogenitas memiliki pengaruh yang paling besar dalam menentukan fungsi diskriminasi (masing-masing -2,117 dan 1,804), sedangkan fitur energi dan max probability memiliki pengaruh yang paling kecil (masing-masing -0,274 dan 0,127).

Selanjutnya analisis diskriminan dilakukan tanpa melibatkan fitur energi dan fitur max probability. Hasilnya ditunjukkan pada tabel 9 – 12 berikut ini.

Tabel 9. Hasil Klasifikasi Tanpa Fitur Energi dan Max Probability (jarak = 1, arah = 0°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results^{b,c}

		Predicted Group Membership		Total
		0	1	
Original	Count	0	6	23
		1	54	55
	%	0	26.1	100.0
		1	98.2	100.0
Cross-validated ^a	Count	0	6	23
		1	54	55
	%	0	26.1	100.0
		1	98.2	100.0

- a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
- b. 91.0% of original grouped cases correctly classified.
- c. 91.0% of cross-validated grouped cases correctly classified.

Tabel 10. Hasil Klasifikasi Tanpa Fitur Energi dan Max Probability (jarak = 1, arah = 45°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results ^{b,c}					
	Kelas	Predicted Group Membership		Total	
		0	1		
Original	Count	0	17	6	23
		1	0	55	55
	%	0	73.9	26.1	100.0
		1	.0	100.0	100.0
Cross-validated ^a	Count	0	17	6	23
		1	1	54	55
	%	0	73.9	26.1	100.0
		1	1.8	98.2	100.0

- a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
 b. 92.3% of original grouped cases correctly classified.
 c. 91.0% of cross-validated grouped cases correctly classified.

Tabel 11. Hasil Klasifikasi Tanpa Fitur Energi dan Max Probability (jarak = 1, arah = 90°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results ^{b,c}					
	Kelas	Predicted Group Membership		Total	
		0	1		
Original	Count	0	17	6	23
		1	0	55	55
	%	0	73.9	26.1	100.0
		1	.0	100.0	100.0
Cross-validated ^a	Count	0	17	6	23
		1	1	54	55
	%	0	73.9	26.1	100.0
		1	1.8	98.2	100.0

- a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
 b. 92.3% of original grouped cases correctly classified.
 c. 91.0% of cross-validated grouped cases correctly classified.

Tabel 12. Hasil Klasifikasi Tanpa Fitur Energi dan Max Probability (jarak = 1, arah = 135°, kelas 0 = kistik, kelas non-kistik =1)

Classification Results ^{b,c}					
	Kelas	Predicted Group Membership		Total	
		0	1		
Original	Count	0	17	6	23
		1	0	55	55
	%	0	73.9	26.1	100.0
		1	.0	100.0	100.0
Cross-validated ^a	Count	0	17	6	23
		1	1	54	55
	%	0	73.9	26.1	100.0
		1	1.8	98.2	100.0

- a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
 b. 92.3% of original grouped cases correctly classified.
 c. 91.0% of cross-validated grouped cases correctly classified.

Berdasarkan data pada tabel 9 – 12, terlihat bahwa tanpa menyertakan fitur energi dan fitur max probability dari analisis fungsi diskriminan, tetap didapatkan kemampuan diskriminasi massa kistik dan massa non-kistik antara 91% - 92,3% .

7. Simpulan dan Saran

Berdasarkan hasil penelitian dan pembahasan diperoleh simpulan bahwa:

- Fitur-fitur berdasar matriks GLCM dapat digunakan untuk membedakan antara massa kistik dan massa non-kistik pada citra ultrasonografi.
- Fitur-fitur penting untuk membedakan citra massa kistik dan massa nonkistik pada citra ultrasonografi adalah: inersia, entropi, homogenitas, kontras, inverse, dan korelasi.

Penelitian selanjutnya akan dilakukan untuk menemukan:

- Berapa ukuran ROI (*region of interest*) yang optimum yang dapat digunakan analisis citra berdasar fitur GLCM?
- Berapa resolusi citra terendah yang dapat digunakan untuk membedakan citra massa kistik dan massa nonkistik pada analisis citra berdasar fitur GLCM?

REFERENCES

- [1] Albregtsen, Fritz, *Statistical Texture Measures Computed from Gray Level Cooccurrence Matrices*, Monograph, Image Processing Laboratory Department of Informatics University of Oslo, 1995 [Online], URL: <http://www.ifi.uio.no/in384/info/glcmm.ps>, (download: February10, 2008).
- [2] Hall-Beyer, Mryka, *The GLCM Tutorial Home Page*, [Online], URL: http://www.fp.ucalgary.ca/mhallbey/the_glcmm.htm, (download: January 7, 2008).
- [3] Hauta-Kasari, Markku, *Computational Techniques for Spectral Image Analysis*, Thesis, Lappeenranta University of Technology Finland, 1999, [Online], URL: ftp://ftp.cs.joensuu.fi/pub/color/theses/dr_thesis_markku_hauta-kasari.pdf, (download tanggal 12 Februari 2008)
- [4] Haralick, R.M. et. al., 1973, *Textural Features for Image Classification*, IEEE Transaction on System, Man, and Cybernetics, Vol. SMC-3.
- [5] Kulak, Eray, 2002, *Analysis of Textural Image Features for Content Based Retrieval*, Thesis, Sabanci University.
- [6] Palmer, P.E.S, 1996, *Panduan Pemeriksaan Diagnostik USG*. Jakarta: Penerbit Buku Kedokteran EGC.