

## Studi Banding antara Metode Minimum Distance dan Gaussian Maximum Likelihood Sebagai Pengklasifikasi Citra Multispektral

Eko Adi Sarwoko

Jurusan Matematika FMIPA UNDIP Semarang  
Jl. Prof. Soedarto, SH Tembalang - Semarang

### Abstrak

Metode *Minimum Distance* merupakan metode sederhana (metrik jarak) yang digunakan untuk alat pengklasifikasi pada masalah pengenalan pola. Demikian pula Metode *Gaussian Maximum Likelihood*, informasi yang ditunjukkan metode ini cukup lengkap, karena selain rata-rata, juga memperhitungkan informasi variansi pada variable yang diukur. Tulisan ini mengkaji mekanisme dan kinerja kedua metode, yang diujikan pada data citra multispektral daerah Riau dan Jawa Tengah. Dari hasil penelitian ini, ternyata penggunaan metode *Gaussian Maximum Likelihood* mempunyai akurasi yang lebih baik dibandingkan metode *Minimum Distance*.

**Kata kunci** : klasifikasi citra, citra multispektral, metode pengklasifikasi *minimum distance*, metode *Gaussian maximum likelihood*, pendekatan geometris, pendekatan statistic, matriks *Population Confusion*, *Average Accuracy*, *Overall Accuracy*.

### 1. PENDAHULUAN

Klasifikasi citra merupakan proses untuk mengelompokkan seluruh piksel pada suatu citra ke dalam sejumlah kelas obyek tertentu, sedemikian sehingga setiap kelas merepresentasikan suatu entitas dengan properti yang spesifik (Chang, 2000). Setiap kelas obyek terdiri dari sekumpulan obyek-obyek yang memiliki deskriptor atau ciri-ciri tertentu. Setiap obyek dalam kelas dapat

direpresentasikan dalam bentuk vektor ciri  $\mathbf{x}$ , dengan  $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ x_n \end{bmatrix}$  dan  $x_1, x_2, \dots,$

$x_n$  menunjukkan ciri dari kelas, serta  $n$  menunjukkan jumlah ciri (Gonzalez, 1993).

Setiap kelas obyek  $k_j$  memiliki suatu vektor pola kelas  $w_j$  yang mencirikan kelas tersebut. Pola kelas  $w_j$  dapat berupa vektor ciri rata-rata,  $m_j = \frac{1}{N_j} \sum_{x \in k_j} x$  dengan  $j=1,2,\dots,M$  serta  $M$  adalah jumlah kelas, dan  $N_j$  adalah jumlah vektor pola dari kelas  $k_j$ .

Untuk citra multispektral, kelas-kelas obyek yang mungkin ada diantaranya adalah kelas sungai, kelas hutan, kelas perkebunan, kelas perumahan, kelas awan, kelas hutan kabut, dsb. Pada pengklasifikasian suatu obyek dengan metode *Minimum Distance* metrik yang digunakan adalah jarak, sedang pada metode *Gaussian Maximum Likelihood* selain rata-rata, juga memperhitungkan variansi variabel yang diukur.

Dalam penelitian terdahulu, dengan kasus citra Saguling pada proporsi sampel pelatihan 30% dan proporsi sample pengujian 70% menunjukkan metode *Minimum Distance* kinerja akurasi klasifikasinya sebesar 91.14% (Sarwoko, 2003). Untuk mengetahui mekanisme dan kinerja kedua metode secara menyeluruh, dilakukan eksperimen pada saat yang bersamaan untuk mengevaluasi hasil klasifikasi, yang diuji cobakan pada citra multispektral daerah Riau dan Jawa Tengah.

## **2. PERBANDINGAN SISTEM PENGAMBILAN KEPUTUSAN**

### **2.1. Pengambilan Keputusan pada Pengklasifikasi Minimum Distance.**

Jika setiap piksel dalam citra direpresentasikan dalam vektor ciri  $u$ , maka salah satu cara untuk menentukan keanggotaan  $u$  adalah dengan memasukkannya ke kelas yang terdekat jaraknya dengan  $u$ . Jarak antara vektor ciri  $u$  dan vektor pola kelas ke- $j$  ( $m_j$ ) dapat dihitung menggunakan jarak *Euclidean* dengan formula

$$D_j(u) = |u - m_j|$$

dengan  $j=1,2,\dots,M$  dan  $|a| = (a^T a)^{1/2}$  adalah *norm Euclidean*. Vektor pola  $u$  dimasukkan ke kelas  $k_j$  jika  $D_j(u)$  merupakan jarak terdekat dibandingkan  $D_p(u)$ ,  $1 < p < M$  dengan  $p \neq j$ . Metode ini dikenal sebagai klasifikasi dengan pendekatan geometrik (Hudson, 1987).

### **2.2. Pengambilan Keputusan pada Pengklasifikasi Gaussian Maximum Likelihood.**

Pada metode Pengklasifikasi *Gaussian Maximum Likelihood*, selain variabel rata-rata juga memperhatikan variansi dari vektor ciri obyek-obyek yang ada pada kelas tersebut. Informasi ini berguna untuk mengetahui sebaran setiap

kelas dari variabel yang diukur. Fungsi Bayes untuk kelas ke  $k_j$  adalah  $d_j(u) = p(u|k_j)P(k_j)$ , dengan  $P(k_j)$  adalah probabilitas vektor ciri  $u$  tergolong dalam kelas  $k_j$ . Jika sebaran vektor ciri setiap kelas diasumsikan memiliki sebaran Gaussian, maka  $p(u|k_j)$  dapat dirumuskan sebagai

$$p(u|k_j) = \frac{1}{(2\pi)^{n/2} |C_j|^{1/2}} \exp \left[ -\frac{1}{2} (u - m_j)^T C_j^{-1} (u - m_j) \right] \quad \dots\dots\dots 1$$

dengan  $C_j$  adalah matriks kovariansi untuk kelas  $k_j$ . Matriks kovariansi ini dirumuskan sebagai  $C_j = \frac{1}{N_j} \sum_{x \in k_j} xx^T - m_j m_j^T$

Dengan adanya bentuk eksponensial ini, akan lebih mudah dengan menghitung  $\ln$  dari fungsi keputusan dibandingkan rumus asli. Sehingga formulanya dapat ditulis sebagai

$$d_j(u) = \ln [p(u|k_j) p(k_j)] = \ln p(u|k_j) + \ln p(k_j) \quad \dots\dots\dots 2$$

Dengan substitusi persamaan (1) ke persamaan (2) diperoleh

$$d_j(u) = \ln p(k_j) - \frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |C_j| - \frac{1}{2} [(u - m_j)^T C_j^{-1} (u - m_j)] \quad \dots\dots\dots 3$$

Nilai  $\frac{n}{2} \ln 2\pi$  sama untuk setiap kelas, sehingga dapat dihilangkan, sehingga persamaan (3) dapat ditulis menjadi

$$d_j(u) = \ln p(k_j) - \frac{1}{2} \ln |C_j| - \frac{1}{2} [(u - m_j)^T C_j^{-1} (u - m_j)] \quad \dots\dots\dots 4$$

Metode ini disebut juga sebagai pengklasifikasi Bayes untuk vektor ciri kelas yang diasumsikan memiliki sebaran Gaussian, vektor ciri  $u$  dimasukkan ke kelas  $k_j$  jika  $d_j(u)$  merupakan probabilitas terbesar dibandingkan  $d_p(u)$ ,  $1 < p < M$  dengan  $p \neq j$ . Metode ini merupakan pendekatan probabilistik, dengan fungsi keputusan yang digunakan adalah fungsi Bayes (Hudson, 1987). Menurut Duda (1973), metode ini mempunyai kinerja baik untuk penginderaan jarak jauh, karena sangat cocok untuk kelas obyek yang distribusi keabuannya mendekati model Gaussian.

### 3. INDIKATOR PERHITUNGAN AKURASI KLASIFIKASI

Untuk mengukur kehandalan metode klasifikasi, dilakukan perhitungan akurasi klasifikasi dengan menggunakan matriks *Population Confusion*, yaitu matriks berordo  $m$ , dengan  $m$  adalah jumlah kelas. Elemen  $a_{ij}$  menunjukkan jumlah piksel citra yang sebenarnya berasal dari kelas  $j$ , namun diklasifikasikan sebagai anggota kelas  $i$ . Sehingga jumlah piksel pada kelas yang benar adalah  $a_{ii}$  (Rennols, 1999). Indikator perhitungan matriks ini adalah :

a. Rata-rata Akurasi setiap Kelas (*Average Accuracy*)

Akurasi klasifikasi pada kelas  $i$  adalah prosentase perbandingan jumlah piksel citra yang benar di kelas  $i$  dengan jumlah keseluruhan piksel (baik piksel yang benar dalam kelas  $i$  ataupun tidak), dapat dirumuskan dengan  $\frac{a_{ii}}{\sum_j a_{ij}}$ .

b. Total akurasi klasifikasi (*Overall Accuracy*)

Total akurasi klasifikasi dirumuskan sebagai prosentase perbandingan jumlah piksel yang berada di kelas yang benar dengan jumlah piksel keseluruhan, yang

dirumuskan sebagai :  $\frac{\sum_i a_{ii}}{\sum_{i,j} a_{ij}}$

### 4. HASIL EKSPERIMEN

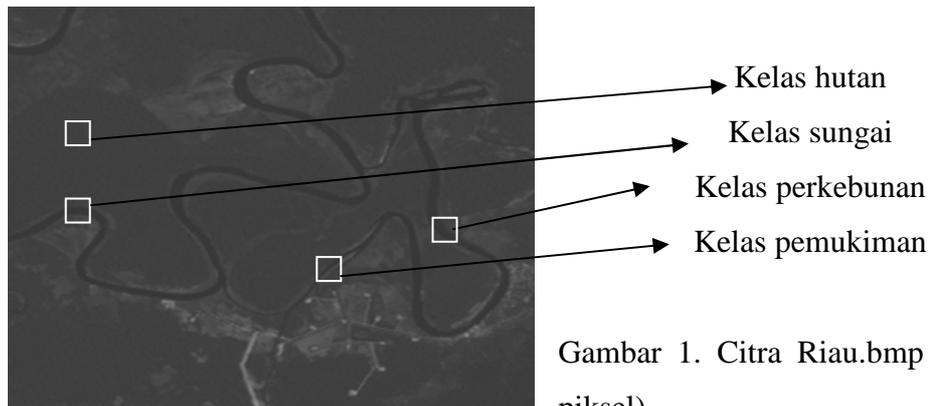
#### 4.1. Citra Masukan

Data citra yang digunakan adalah citra multispektral daerah Riau (enam band : band 1 sampai band 6) dan citra Lanier.bmp (sumber Bakosurtanal RI) band 1-6. Proporsi sampel pada eksperimen ini, diambil dengan proporsi pelatihan 30% dan pengujian 70%.

a. Citra Riau.bmp

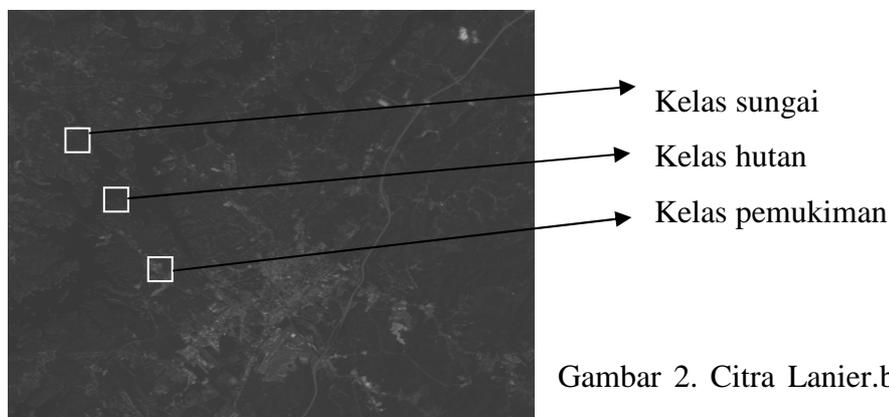
Citra optik ini adalah citra Riau.bmp (sumber Lapan RI) band 1-6, contoh lihat gambar 1, merupakan daerah yang diasumsikan memiliki 4 kelas, yaitu sungai, hutan, perkebunan dan pemukiman. Citra ini berukuran 512x512 piksel,

kelas 1 diambil dari 8 lokasi terpisah (200 piksel), 12 lokasi kelas 2 (300 piksel), 8 lokasi kelas 3 (200 piksel) dan 4 lokasi kelas 4 (100 piksel), yang masing-masing sampel piksel berukuran 5x5.



Gambar 1. Citra Riau.bmp (512x512 piksel).

b. Citra Lanier.bmp

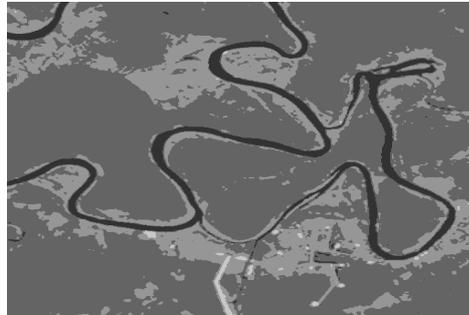


Gambar 2. Citra Lanier.bmp (512x512 piksel).

Citra optik yang kedua adalah citra Lanier.bmp (sumber Bakosurtanal RI) band 1-6. Citra Lanier ini, lihat gambar 4.4., merupakan citra daerah Jawa Tengah yang diasumsikan memiliki 3 kelas, yaitu kelas 1 sungai, kelas 2 hutan, dan kelas 3 pemukiman. Citra ini berukuran 512x512 piksel, sampel kelas berukuran 150 piksel masing-masing berukuran 5x5 piksel.

## 4.2. Citra Keluaran

a. Hasil Uji Coba Klasifikasi pada Citra Riau.bmp



Gambar 3. Citra Riau hasil klasifikasi dengan *Minimum Distance* (95%)

Tabel 1. *Confusion Matrix* pelatihan

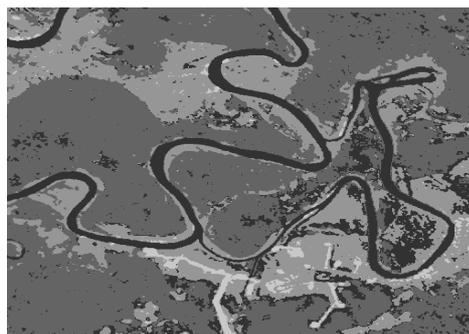
Klasifikasi *Minimum Distance*

	1	2	3	4
1	54	0	0	0
2	6	90	0	0
3	0	0	60	2
4	0	0	0	28

Tabel 2. *Confusion Matrix* Pengujian

Klasifikasi *Minimum Distance*

	1	2	3	4
1	120	0	0	0
2	20	210	1	0
3	0	0	139	7
4	0	0	0	63



Gambar 4. Citra Riau hasil klasifikasi dengan *Maximum Likelihood* (94.11%)

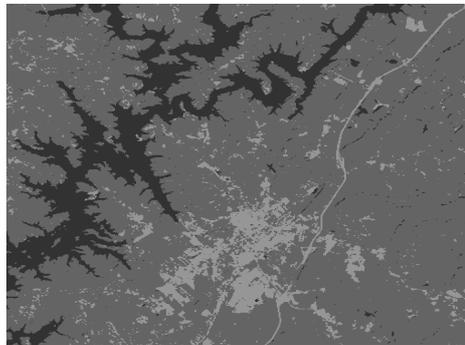
Tabel 3. *Confusion Matrix* pelatihan  
 Klasifikasi *Maximum Likelihood*

	1	2	3	4
1	53	0	0	0
2	7	89	0	0
3	0	1	60	3
4	0	0	0	27

Tabel 4. *Confusion Matrix* Pengujian  
 Klasifikasi *Maximum Likelihood*

	1	2	3	4
1	119	0	0	0
2	21	210	2	0
3	0	0	138	10
4	0	0	0	60

b. Hasil Uji Coba Klasifikasi pada Citra Lanier.bmp



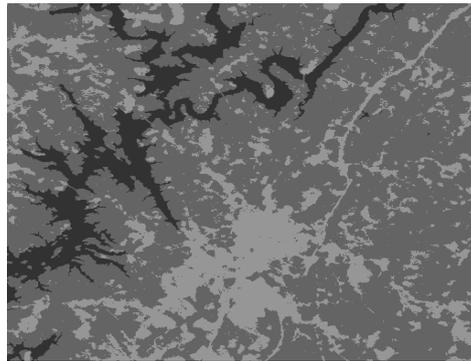
Gambar 5. Citra Lanier Hasil Klasifikasi dengan *Minimum Distance* (97%)

Tabel 5. *Confusion Matrix* pelatihan  
 Klasifikasi *Minimum Distance*

	1	2	3
1	50	0	0
2	0	49	2
3	0	1	48

Tabel 6. *Confusion Matrix* Pengujian  
 Klasifikasi *Minimum Distance*

	1	2	3
1	98	0	0
2	2	99	6
3	0	1	94



Gambar 6. Citra Lanier hasil klasifikasi dengan *Maximum Likelihood* (97.67%)

Tabel 7. *Confusion Matrix* pelatihan

Klasifikasi *Maximum Likelihood*

	1	2	3
1	50	0	0
2	0	50	0
3	0	0	50

Tabel 8. *Confusion Matrix* Pengujian

Klasifikasi *Maximum Likelihood*

	1	2	3
1	93	0	0
2	7	100	0
3	0	0	100

#### 4.3. Evaluasi Dan Analisis

Dari hasil uji coba ini secara umum diperlihatkan bahwa citra hasil menunjukkan perbaikan kualitas citra. Selain itu diperoleh suatu gambaran bahwa dengan proporsi sampel pelatihan 30% dan pengujian 70% menunjukkan hasil akurasi klasifikasi pada citra multispektral Riau dengan metode *Minimum Distance* dan metode *Gaussian Maximum Likelihood* masing masing adalah 95% dan 94,11%. Sedang pada citra multispektral Lanier (Jawa Tengah) dengan metode *Minimum Distance* dan metode *Gaussian Maximum Likelihood* masing masing adalah 97% dan 97,67%.

#### 4. KESIMPULAN

Berdasarkan penelitian ini, secara umum dapat disimpulkan bahwa sistem pengambilan keputusan kedua metode berbeda. Pada metode *Minimum Distance* merupakan pendekatan geometris (jarak), sedang pada pengklasifikasi metode *Gaussian Maximum Likelihood* menggunakan pendekatan statistik (rata-rata, variansi).

Secara teori pendekatan metode pengklasifikasi metode *Gaussian Maximum Likelihood* haruslah lebih baik dibandingkan dengan metode *Minimum Distance*. Hal ini disebabkan pada *Minimum Distance* hanya memperhitungkan aspek jarak saja, sedang pada metode *Gaussian Maximum Likelihood* memperhitungkan informasi sekitar seperti rata-rata, variansi dan sebarannya. Namun apabila dilihat pada hasil eksperimen kasus citra Riau, ternyata metode *Gaussian Maximum Likelihood* (94,11) akurasi tidak lebih baik dibandingkan dengan metode *Minimum Distance* (95%). Hal ini kemungkinan disebabkan informasi yang diperoleh pada sampel tidak berdistribusi normal, sehingga variansinya menjadi sangat besar. Hal ini akan sangat mempengaruhi kinerja klasifikasi. Namun secara umum kinerja klasifikasi metode *Gaussian Maximum Likelihood* lebih baik dibandingkan dengan metode *Minimum Distance*.

## **5. DAFTAR PUSTAKA**

- Chang, C.I., H. Ren, *An Experiment-Based Quantitative and Comparative Analysis of Target Detection and Image Classification Algorithms for Hyperspectral Imagery*, IEEE Trans. On Geoscience and Remote Sensing, vol.38, no.2, Maret 2000
- Duda, R.O., P.E. Hart, *Pattern Classification and Scene Analysis*, New York : Wiley, 1973
- Gonzalez, Rafael C, Richard E.Woods, *Digital Image Processing*, Addison-Wesley Publishing Company, Inc., 1993
- Hudson, William D., *Digital Classification of Landsat Multispectral Scanner Data – An Introduction*, Research Report Natural Resources From The Michigan State University Agricultural Experiment Station East Lansing, No 483, Agustus 1987
- Lillesand, T.M., R.W. Kiefer, *Remote Sensing and Image Interpretation*, John Wiley & Sons, Inc., 1994
- Rennols, K., *Image Classification Accuracy from LANDSAT Data for Tropical Rain Forest in Sumatra : An Evaluation of ERDAS image software*,

*Data Management and Modelling Using Remote Sensing and GIS for Tropical Forest Land Inventory*, Rodeo International Publishers, Jakarta, 1999

Richards, J.A., *Remote Sensing Digital Image Analysis, An Introduction*, Springer-Verlag Berlin Heidelberg, 1986

Sarwoko, E.A., *Analisis Pengklasifikasi Minimum Distance pada Penginderaan Jarak Jauh, Studi Kasus Citra Saguling*, Prosiding Seminar Nasional Matematika dan Statistika VI Kerjasama Jurusan Statistika ITS Surabaya dengan Himpunan Alumni Pascasarjana UGM Yogyakarta, Surabaya, 11 Oktober 2003.