Bab 3
Pendekatan Statistik
Klasifikasi Citra Penginderaan Jauh

Klasifikasi citra penginderaan jarak jauh (inderaja) merupakan proses penentuan piksel-piksel masuk ke dalam suatu kelas obyek tertentu. Pendekatan *klasifikasi citra* yang dapat digunakan antara lain pendekatan statistik. Pendekatan statistik yang telah sering digunakan antara lain *Euclidean Distance (ED)*, *Mahalanobis Distance (MD)*, *Maximum Likelihood (ML)/Bayesian* [2]. metode *euclidean distance* menghitung jarak spektrum setiap elemen citra terhadap nilai rerata vektor dari setiap kelas. Apabila jarak spektrum telah dihitung, maka suatu elemen (piksel) akan masuk ke dalam suatu kelas yang jaraknya paling dekat. Metode *mahalanobis distance* memiliki kesamaan dengan *euclidean distance*, namun pada perhitungan jaraknya menggunakan matriks *kovarian*. Metode *maximum likelihood/Bayesian* mengasumsikan bahwa setiap elemen citra yang memiliki suatu nilai probabilitas *posterior* sama tergolong dalam suatu kelas yang sama. Metode yang terakhir ini adalah metode yang paling baik dibanding metode sebelumnya untuk model sebaran data normal [2]. Ketiga metode dari pendekatan statistik tersebut memerlukan model sebaran normal yang sangat bergantung pada parameter *median* dan *standar deviasi*.

2.1 Euclidean Distance (ED)

Metode *Euclidean Distance* (ED) menghitung jarak spektrum setiap piksel terhadap rerata vektor setiap *signatur* atau ciri [2, 19]. Sebuah piksel dimasukkan pada suatu kelas yang jaraknya paling dekat dengan signaturnya. Keuntungan metode ini adalah semua piksel akan terklasifikasi, dan merupakan metode yang paling cepat secara komputasi. Kelemahannya karena tidak memasukkan perhitungan *variabilitas* suatu kelas, sehingga akan ada nilai piksel yang cenderung jauh dari rerata signatur kelasnya. Contoh kelas liputan lahan perkotaan yang terdiri dari piksel-piksel dengan *variansi* tinggi, maka dengan aturan tersebut, terdapat kemungkinan adanya piksel yang terklasifikasi dengan tidak benar. Sebaliknya suatu kelas dengan *variansi* rendah seperti air, cenderung terjadi klasifikasi berlebih *(overclassification)* karena piksel-piksel dari kelas yang bervariansi tinggi ada kemungkinan jaraknya lebih dekat dengan rerata *signatur* kelas air dibanding dengan rerata *signatur* kelas yang sebenarnya.

2.2 Mahalanobis Distance (MD)

Metode *Mahalanobis Distance* (MD) mirip dengan *euclidean distance*, tetapi menggunakan matriks *kovarians* dalam perhitungan jarak tersebut [2, 19]. *Varians* dan *kovarians* digunakan agar piksel dari kelas yang bervariansi tinggi dapat terklasifikasi secara benar walau jarak *euclideannya* dengan rerata *signatur* kelas lain lebih dekat dibanding dengan jarak *euclidean* ke kelasnya. Metode ini dapat mengatasi kelemahan dari metode *euclidean distance*.

$$D = (X - M_c)^T (Cov_c^{-1}) (X - M_c) \quad \ldots\ldots\ldots\ldots\ldots(3.1)$$

Keterangan :

D	:	*jarak mahalanobis,*
X	:	*vektor suatu piksel,*
M_c	:	*vektor rerata signatur kelas C,*
Cov_c	:	*matriks kovarian dalam kelas C.*

Kelemahan metode ini diantaranya adalah menggunakan parameter statistik dengan sebaran data normal, sehingga sangat bergantung pada informasi *prior* dan model sebaran data. Selain itu, metode ini sulit untuk mengklasifikasi piksel-piksel dari kelas yang tidak memiliki batas yang jelas dengan kelas lainnya seperti daerah perubahan dari satu kelas ke kelas lain yang disebut *zone perubahan*.

2.3 Maximum Likelihood (ML)

Metode *Maximum Likelihood (ML)* didasarkan pada asumsi bahwa setiap piksel memiliki *probabilitas* (kemiripan) yang berbeda untuk setiap kelas. Suatu piksel akan dimasukkan ke dalam suatu kelas bila kemiripannya paling besar dibanding terhadap kelas lainnya [2, 20]. Metode *maximum likelihood* tidak saja memperhatikan informasi rerata vektor kelas, akan tetapi juga memperhatikan informasi variansi vektor kelas tersebut [2]. Fungsi keputusan untuk kelas w_j dinyatakan :

$$d_j(x) = p(x|\,\omega_j)p(\omega_j) \qquad (3.2)$$

Keterangan*:*

$p(\omega_j)$: *probabilitas prior vektor x untuk kelas* w_j

$p(x|\,\omega_j)$: *probabilitas bersyarat (probability density function/pdf)*

Dengan sumsi sebaran Gaussian, maka $p(x|\,\omega_j)$ dapat dirumuskan sebagai berikut:

$$p(x\,|\,\omega_j) = \frac{1}{(2\pi)^{\pi/2}\,|Cj|^{1/2}}\,exp\left[-\frac{1}{2}(x-m_j)^T C_j^{-1}(x-m_j)\right] \qquad (3.3)$$

$$C_j = \frac{1}{N_j}\sum_{x\in w_j} x\,x^T - m_j\,m_j^T \qquad (3.4)$$

Keterangan :

C_j : *matriks kovarians untuk kelas* w_j.

j : *nomor kelas*

m : *rerata sampel*

Bentuk sebaran Gaussian lebih mudah apabila perhitungan menggunakan *logaritma natural* maka fungsi keputusan menjadi sebagai berikut:

$$d_j(x) = ln\ [p(x|\omega_j)\ p(\omega_j)] = ln\ p(x|\omega_j) + ln\ p(\omega_j) \quad \ldots\ldots\ldots\ldots\ldots(3.5)$$

Substitusi (3.3) kedalam (3.5) diperoleh fungsi keputusan sebagai berikut:

$$d_j(x) = ln\ p(\omega_j) - \frac{\pi}{2} ln\, 2\pi - \frac{1}{2} ln\,|C_j| - \frac{1}{2}[(x-m_j)^T C_j^{-1}(x-m_j)] \quad \ldots\ldots\ldots(3.6)$$

Nilai $\frac{\pi}{2} ln(2\pi)$ berharga sama untuk semua kelas, maka dapat dihilangkan sehingga persamaan (3.6) menjadi sebagai berikut :

$$d_j(x) = ln\ p(\omega j) - \frac{1}{2} ln\,|Cj| - \frac{1}{2}[(x-m_j)^T C_j^{-1}(x-m_j)] \quad \ldots\ldots\ldots\ldots..(3.7)$$

Berdasarkan kaidah keputusan tersebut, vektor x dimasukkan ke kelas w_j jika $d_j(x)$ merupakan probabilitas terbesar dibanding $d_f(u)$, $1 \leq f \leq M$, $f \neq j$.

Metode ML berpotensi menghasilkan kesalahan klasifikasi seperti pada kelas hutan dan bukan hutan. Metode ini memberikan hasil bahwa walaupun tidak terdapat hutan yang salah klasifikasi, tetapi terdapat area bukan hutan yang diklasifikasi sebagai hutan. Hal ini dikarenakan terdapat bukan hutan yang memiliki kemiripan yang lebih besar dengan kelas hutan dibanding terhadap kelas bukan hutan. Untuk mengatasi masalah ini diperlukan pengaturan sehingga walaupun memiliki probabilitas lebih rendah, namun masih bisa dikategorikan dalam kelas yang benar. Pengaturan tersebut menggunakan probabilitas *prior* yang berbeda untuk setiap kelas [19]. Para peneliti yang telah mempelajari pendekatan statistik antara lain Benediktsson, dkk. (1989), Lee (1990), Murni (1996), dan Schistad dkk. (1997).