PENGEMBANGAN APLIKASI *MUNSELL SOIL COLOR DETECTION CHART INDEX* MENGGUNAKAN METODE *SUPPORT VECTOR MACHINE*

Mungki Astiningrum[1], Putra Prima Arhandi[2], Elly Fatmawati[3]

[1,2]Program Studi Teknik Informatika, Jurusan Teknologi Informasi, [3]Politeknik Negeri Malang
[1] mungki.polinema@gmail.com , [2] putraprima@polinema.ac.id, [3]ellyfatmawati95@gmail.com

Abstrak

Salah satu sifat tanah yang paling sering digunakan oleh para peneliti untuk menggambarkan dan mengklasifikasikan tanah adalah warna. Warna merupakan sifat fisik yang dapat memberikan informasi tentang beberapa karakteristik paling penting dari suatu tanah seperti, komposisi mineral, usia, dan proses pembentukan. Dengan mengetahui warna dari suatu tanah, maka dapat ditentukan kegunaan tanah tersebut apakah cocok untuk lahan pertanian, perkebunan, perumahan, dan lain sebagainya, disesuaikan dengan kondisi tanahnya. Dalam mengidentifikasi warna tanah, para peneliti menggunakan standar warna pada buku *Munsell Soil Color Chart*. Banyaknya variasi warna pada buku ini menyebabkan peneliti membutuhkan kecermatan dan waktu lama dalam penentuan warna tanah. Keterbatasan indera penglihatan manusia juga mampu menjadi faktor terjadinya kesalahan sehingga informasi yang diperoleh menjadi kurang akurat. Oleh karena itu, dikembangkanlah aplikasi menggunakan pengolahan citra digital untuk mengidentifikasi warna tanah berdasarkan *Munsell Soil Color Chart* menggunakan metode *Support Vector Machine* (SVM) dan ekstraksi fitur warna Modus dan *Mean*. Proses diawali dengan melakukan *cropping* terhadap citra tanah inputan menjadi 170x200, kemudian mengurangi *noise* pada citra masukan menggunakan median filter dan mengekstraksi fitur dengan Modus dan *Mean*. Selanjutnya proses *training* dan *testing* data menggunakan metode SVM. Pengujian yang dilakukan menghasilkan tingkat keakuratan mencapai 93,33% untuk masing – masing ekstraksi.

Kata kunci: citra digital, deteksi warna tanah, *Munsell Soil Color Chart,* tanah

1. Pendahuluan

1.1 Latar Belakang

Tanah adalah bangunan alam yang tersusun atas horison-horison yang terdiri atas bahan mineral dan organik, biasanya tak padu, mempunyai tebal yang berbeda-beda dan yang berbeda pula dengan bahan induk yang ada di bawahnya dalam hal morfologi, sifat dan susunan fisik, sifat dan susunan kimia, dan sifat-sifat biologi Sutanto (2005). Salah satu sifat yang paling sering digunakan oleh para peneliti untuk menggambarkan dan mengklasifikasikan tanah adalah warna. Warna merupakan sifat fisik yang dapat memberikan informasi tentang beberapa karakteristik paling penting dari suatu tanah seperti, komposisi mineral, usia, dan proses pembentukan Jordan (2014). Hasil klasifikasi tanah selanjutnya dijadikan dasar penilaian kesesuaian lahan untuk pertanian, perhutanan, dan sebagainya.

Dalam penelitian sumberdaya tanah saat ini, peneliti menggunakan *Munsell Soil Color Chart* (MSCC) sebagai acuan penentuan warna tanah. Cara menggunakan MSCC ini adalah dengan membandingkan kemiripan warna tanah asli dengan pedoman warna yang ada dalam buku tersebut. Banyaknya variasi warna dalam buku ini mencapai lebih dari 276 macam warna membuat peneliti membutuhkan waktu yang lama dalam penentuan warna tanah. Belum lagi karena keterbatasan indera penglihatan manusia dalam pengenalan warna memungkinkan terjadinya kesalahan sehingga informasi yang diperoleh kurang akurat. Berdasarkan hal – hal tersebut, maka pada penelitian ini dibuatlah pengembangan suatu aplikasi menggunakan pengolahan citra digial untuk mengidentifikasi warna tanah berdasarkan *Munsell Soil Color Chart* menggunakan metode *Support Vector Machine* sebagai metode untuk pengklasifikasian warna tanah dan ekstraksi fitur warna Modus dan *Mean* dengan harapan bahwa aplikasi pada penelitian ini akan memberikan alternatif kepada peneliti dan pengguna lainnya untuk mempermudah dan mempercepat penentuan warna tanah dengan hasil yang akurat.

1.2 Rumusan Masalah

a. Bagaimana mengembangkan suatu aplikasi yang berfungsi untuk mendeteksi warna tanah

menggunakan standar warna pada *Munsell Soil Color Chart* agar mempermudah dan mempercepat dalam penentuan warna tanah?

b. Bagaimana tingkat akurasi metode *Support Vector Machine* (SVM) dalam melakukan klasifikasi warna tanah?

1.3 Batasan Masalah

1. Data *chart* warna yang digunakan adalah 5 index warna yang diambil secara acak pada buku Munsell *Soil Color Chart*.
2. Warna tanah yang diklasifikasikan adalah warna *top soil*.
3. Informasi yang ditampilkan berupa klasifikasi index warna tanah yang diuji sesuai dengan pedoman Munsell *Soil Color Chart*.

2. Landasan Teori

2.1 Tanah

Salah satu komponen dari lahan yang merupakan komponen utama adalah tanah Hardjowigeno (1992). Tanah adalah kumpulan dari benda alam di permukaan bumi yang tersusun dalam horison-horison, terdiri dari campuran bahan mineral, bahan organik, air dan udara, dan merupakan media untuk tumbuhnya tanaman Hardjowigeno (1987).

2.2 Warna Tanah

Warna tanah berfungsi sebagai penunjuk dari sifat tanah, karena warna tanah dipengaruhi oleh beberapa faktor yang terdapat dalam tanah tersebut. Saat ini, untuk menentukan warna tanah peneliti membandingkan warna pada buku Munsell Soil Color Chart. Diagram warna baku Munsell disusun dalam tiga variable Madjid (2007) yaitu:

a. *Hue* merupakan warna spektrum yang dominan sesuai dengan panjang gelombangnya.
b. *Value* menunjukkan gelap terangnya warna sesuai dengan banyaknya sinar yang dipantulkan.
c. *Chroma* menunjukkan kemurnian atau kekuatan dari warna spektrum.

2.3 Citra Digital

Citra atau gambar dapat didefinisikan sebagai sebuah fungsi dua dimensi f(x,y), di mana x dan y adalah koordinat bidang datar, dan harga fungsi f di setiap pasangan koordinat (x,y) disebut intensitas atau level keabuan (*grey level*) dari gambar di titik itu Hermawati (2013).

2.4 Citra Berwarna RGB

Red (Merah), *Green* (Hijau) dan *Blue* (Biru) merupakan warna dasar yang dapat diterima oleh mata manusia. Setiap *pixel* pada citra warna mewakili warna yang merupakan kombinasi dari ketiga warna dasar RGB.). RGB didasarkan pada teori bahwa mata manusia peka terhadap panjang gelombang 630nm (merah), 530 nm (hijau), dan 450 nm (biru) Rizal (2014).

2.5 Modus RGB

Modus adalah nilai warna yang *pixel* yang paling sering muncul dalam suatu gambar Robbani (2016).

2.6 *Mean* RGB

Mean atau rata-rata adalah penjumlahan semua nilai yang terdapat pada masing-masing komponen R, G, dan B. Kemudian hasil dari penjumlahan tersebut dibagi dengan jumlah *pixel* yang terdapat dalam citra.

$$\mu = \frac{x_1 + x_2 + \cdots + x_n}{N} \quad (1)$$

Keterangan:

μ : Rata-rata
x_n : Nilai *pixel*
N : Jumlah keseluruhan *pixel*

2.7 Median Filter

Median filter merupakan *order-statistics* filter yang paling banyak diketahui, yang mana sesuai namanya, mengganti nilai *pixel* dengan median dari level intensitas dalam *neighbourhood* dari *pixel* itu Woods, dkk (2008).

$$f(x,y) = \underset{(s,t) \in S_{xy}}{median}\{g(s,t)\} \quad (2)$$

2.8 Normalisasi

Normalisasi diperlukan untuk memperkecil *range* nilai dari data agar tidak terlalu besar. Normalisasi mempunyai banyak jenis salah satunya Normalisasi Data Sigmoid. Normalisasi ini memperkecil *range* nilai menjadi 0,1 hingga 0,9 Kusuma (2016).

$$z' = \frac{0.8\ (x-a)}{b-a} + 0{,}1 \quad (3)$$

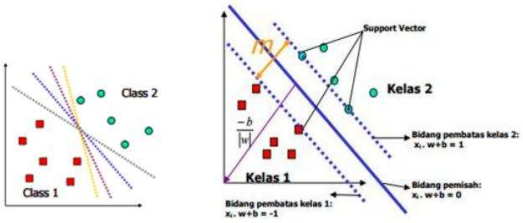
Keterangan:

z : nilai baru setelah dinormalisasi
x : data
b : nilai maximum dari data
a : nilai minimum dari data

2.9 *Support Vector Machine*

Support Vector Machine (SVM) merupakan suatu metode pembelajaran untuk mesin yang baru berdasarkan teori pembelajaran statistik yang diajukan oleh Vapnik di tahun 90an. Inti dari proses pelatihan pada SVM ini adalah usaha untuk

mencari lokasi *hyperplane*. *Hyperplane* pemisah terbaik antara kedua kelas tersebut dapat ditemukan dengan mengukur margin *hyperplane* dan mencari titik maksimalnya. Margin merupakan jarak antara *hyperplane* tersebut dengan data terdekat dari masing-masing kelas. Data yang paling dekat inilah yang disebut *support vector* Kusuma (2016). Jika titik terpisah secara linear fungsi untuk permukaan ini ditentukan oleh persamaan 4.



Gambar 1. Ilustrasi Pemisahan Kelas SVM

$$f(x) = sign\left(\sum_{i=1}^{s} a_i y_i (x_i . x) + b\right)$$
$$(x_i, y_i) \in R^N x \{-1, 1\} \quad (4)$$

Dimana:

a_i = *Langrange Multiplier*
b = bias

Jika kelas tidak terpisah secara linear maka fungsi untuk permukaan ditentukan oleh persamaan 5.

$$f(x) = sign\left(\sum_{i=1}^{s} a_i y_i k(x, y) + b\right) \quad (5)$$

Dimana:

a_i = *Langrange Multiplier*
b = bias
$k(x, y)$ = Fungsi Kernel

Terdapat beberapa fungsi kernel yang digunakan untuk menyelesaikan masalah pada SVM *non-linear* dapat dilihat pada persamaan 6 hingga 8.

Fungsi Kernel Linear:

$$k(x, y) = x . y \quad (6)$$

Fungsi Kernel Polynomial:

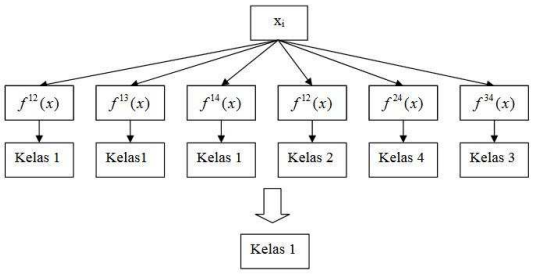
$$k(x, y) = (1 + x . y)^q, q = 1, 2, \ldots, N \quad (7)$$

Fungsi Kernel Radian Basis

$$k(x, y) = exp \frac{\left(-\|x - y\|^2\right)}{2\sigma^2} \quad (8)$$

2.10 *Multiclass* SVM

Metode yang umum digunakan untuk mengimplementasikan *multi class* SVM dengan pendekatan yang pertama adalah Metode *One-Against-All* dan Metode *One-Against-One*. Menggunakan Metode *One-Against-One*, mengharuskan setiap model klasifikasi dilatih pada data dari dua kelas. Terdapat beberapa metode untuk melakukan pengujian setelah keseluruhan model klasifikasi selesai dibangun. Salah satunya adalah metode *voting* Sembiring (2017).



Gambar 2. Klasifikasi dengan *One-Against-One*

Kelas dari data x akan ditentukan dari jumlah suara terbanyak. Jika terdapat dua buah kelas yang jumlah suaranya sama, maka kelas yang indeksnya lebih kecil dinyatakan sebagai kelas dari data. Jadi pada pendekatan ini terdapat $k(k-1)/2$ buah permasalahan *quadratic programming* yang masing – masing memiliki $2n/k$ variabel (n adalah jumlah data pelatihan).

2.11 Metode *Sequential Training* SVM

Metode *Sequential Training* merupakan metode yang digunakan untuk training data agar menghasilkan *hyperplane* yang optimal Kusuma (2016).

1. Melakukan perhitungan kernel dan melakukan inisialisasi parameter – parameter SVM seperti contoh nilai $a_i = 0$; ε = 0,001; γ = 0,01; λ = 1; C = 1; dan nilai Iterasi maksimum = 3.
2. Menghitung matriks Hessian

$$D_{ij} = y_i y_j (K(x_i, x_j) + \lambda^2) \quad (9)$$

3. Melakukan iterasi untuk setiap iterasi = 1, 2, 3, …, n

   Menghitung nilai Ei

$$E_i = \sum_{j=1}^{n} a_i D_{ij} \quad (10)$$

   Menghitung nilai γ dan δa_i

$$\gamma = \frac{konstanta}{\max D_{ij}} = 0,008 \quad (11)$$

$$\delta a_i = min[max[\gamma(1 - E_i), a_i], C - a_i] \quad (12)$$

   Memperbarui nilai a_i

$$a_i = a_i + \delta a \quad (13)$$

4. Proses 1, 2, dan 3 diulangi hingga nilai δα mencapai konvergen ($|\delta\alpha| < \varepsilon$) dan atau ketika nilai iterasi sudah mencapai nilai maksimum, maka iterasi dihentikan.
5. Menghitung nilai $w.x^+$ dan $w.x^-$ untuk mendapatkan nilai bias.

$$w.x^+ = \alpha_i.y_i.K(x.x^-) \quad (14)$$

$$w.x^- = \alpha_i.y_i.K(x.x^-) \quad (15)$$

$$b = -\frac{1}{2}(w.x^+ + w.x^-) \quad (16)$$

Keterangan:

$K(x,x^+)$: Nilai kernel data x dengan data x kelas positif yang memiliki nilai α tertinggi.

$K(x,x^-)$: Nilai kernel data x dengan data x kelas negatif yang memiliki nilai α tertinggi.

b : nilai bias

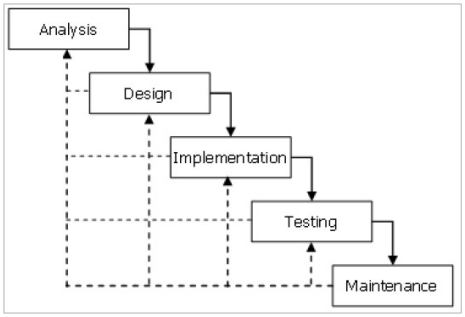
2.12 Akurasi

Akurasi merupakan nilai yang menunjukkan seberapa besar ketepatan yang didapatkan dari sebuah *testing* klasifikasi. Cara menghitung akurasi dapat menggunakan rumus:

$$Nilai\ akurasi = \frac{\Sigma\ hasil\ prediksi\ benar}{\Sigma\ data\ uji\ keseluruhan} \times 100\% \quad (17)$$

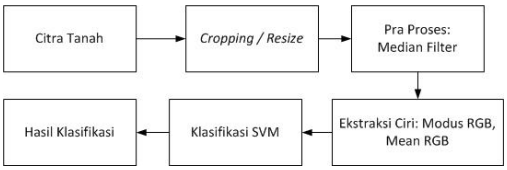
3. Metodologi

3.1 Desain Penelitian



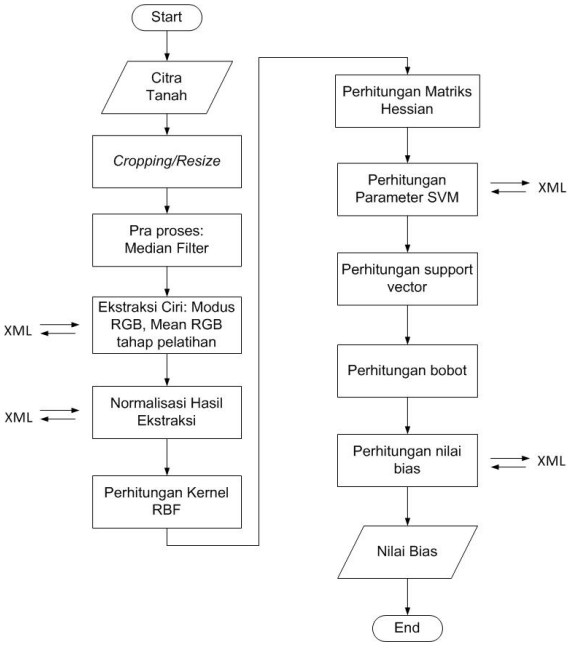
Gambar 3. Diagram Blok Penelitian

4. Perancangan

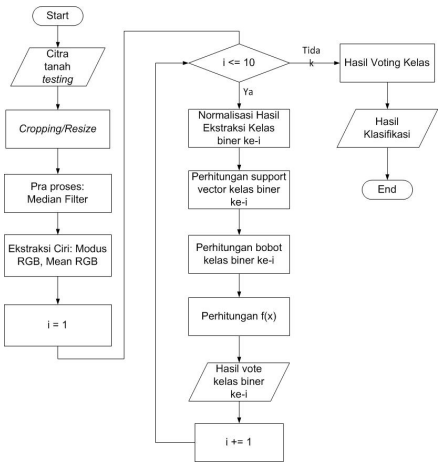


Gambar 4. Blok Diagram Alur Sistem

Data masukan untuk penelitian ini adalah citra digital dengan keluaran berupa hasil klasifikasi. Proses dimulai dengan melakukan *cropping* / *resize* pada citra tanah input menjadi 170x200, apabila citra tersebut belum berukuran 170x200. Namun karena penulis menggunakan data citra yang telah berukuran 170x200, maka proses *cropping* dan *resize* tidak perlu dilakukan. Kemudian melakukan proses pengurangan *noise* pada citra inputan menggunakan median filter. Setelah didapatkan citra tanpa *noise*, dilakukan proses perhitungan Modus dan *Mean* RGB untuk mendapatkan ciri dari warna. Kemudian melakukan proses perhitungan normalisasi data untuk mendapatkan data dalam *range* yang tidak terlalu besar. Proses berikutnya melakukan perhitungan klasifikasi menggunakan metode SVM untuk mendapatkan hasil klasifikasi.



Gambar 5. *Flowchart Training*



Gambar 6. *Flowchart Testing*

5. Uji coba dan Pembahasan

5.1 Pengujian Tingkat Akurasi

Pengujian tingkat akurasi digunakan untuk menguji seberapa besar sistem deteksi warna tanah dengan menerapkan metode *Support Vector Machine* ini berhasil. Pada pengujian ini penulis menggunakan sistem *offline* (tidak *realtime*) dengan memanfaatkan citra tanah yang telah diperoleh oleh penulis sebelumnya.

Pada pengujian ini menggunakan sebanyak 15 sampel tanah yang terdiri dari index warna 5 Y 3/1, 10 YR 3/2, 10 YR 5/4, 2.5 YR 4/2, dan 7.5 YR 5/6 yang diekstraksi menggunakan fitur modus dan *mean*. Berikut merupakan data sampel *testing* beserta hasil pengujian deteksi warna tanah dengan menerapkan metode *Support Vector Machine*.

Tabel. Data Sampel *Testing*

Jumlah Data Tanah	Index Warna
3 buah	5 Y 3/1
3 buah	10 YR 3/2
3 buah	10 YR 5/4
3 buah	2.5 YR 4/2
3 buah	7.5 YR 5/6

Pengujian dilakukan menggunakan parameter lamda = 0,5, gamma = 0,01, complexity = 1, epsilon = 0,001, dan iterasi maksimum = 3.

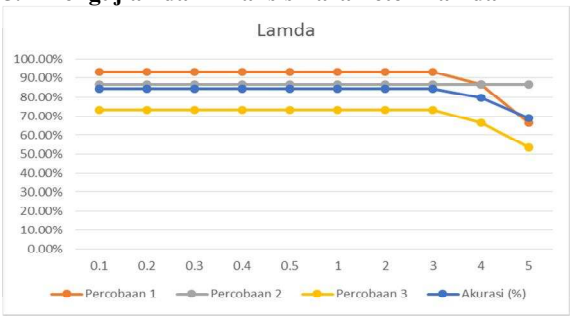
Tabel 2. Hasil Tingkat Akurasi Fitur Modus

Index Warna	Jumlah Sampel	Sesuai	Tidak Sesuai	Tingkat Akurasi
5 Y 3/1	3	3	0	100 %
10 YR 3/2	3	3	0	100 %
10 YR 5/4	3	3	0	100 %
2.5 YR 4/2	3	2	1	66,67 %
7.5 YR 5/6	3	3	0	100 %
Rata – rata akurasi				93,33 %

Tabel 3. Hasil Tingkat Akurasi Fitur *Mean*

Index Warna	Jumlah Sampel	Sesuai	Tidak Sesuai	Tingkat Akurasi
5 Y 3/1	3	3	0	100 %
10 YR 3/2	3	3	0	100 %
10 YR 5/4	3	3	0	100 %
2.5 YR 4/2	3	2	1	66,67 %
7.5 YR 5/6	3	3	0	100 %
Rata – rata akurasi				93,33 %

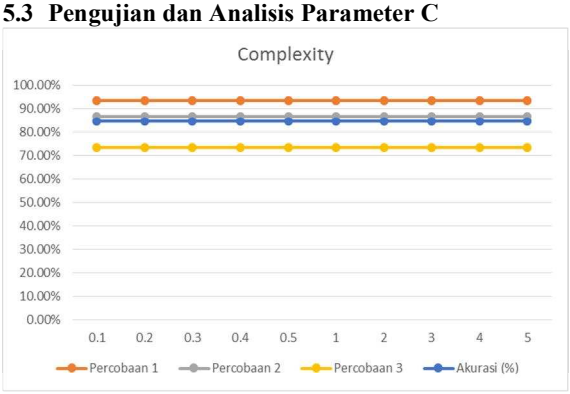
5.2 Pengujian dan Analisis Parameter Lamda



Gambar 7. Grafik Hasil Uji Parameter Lamda

Pada grafik pengujian dapat dilihat bahwa semakin besar *lamda* maka akurasi yang diperoleh akan semakin menurun. Hal ini karena semakin besar *lamda,* maka semakin besar pula kemungkinan terjadinya *overfitting* (atau disebut juga *high variance*) yang disebabkan oleh fungsi hipotesis yang sesuai (fit) pada hampir semua data dari *training* set tetapi tidak sesuai untuk data secara umum pada proses prediksi data baru selanjutnya.

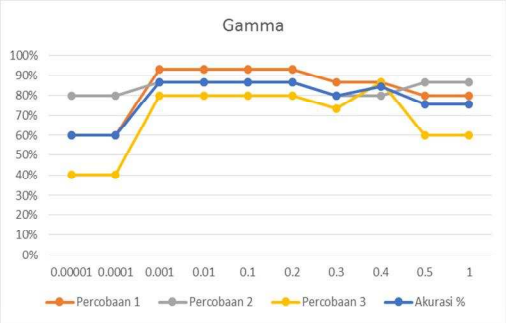
5.3 Pengujian dan Analisis Parameter C



Gambar 8. Grafik Hasil Uji Parameter C

Dari Gambar 8 diperoleh hasil nilai C yang stabil dan tidak terjadi perubahan. Dari hal ini dapat diketahui bahwa mengubah nilai C dapat menghasilkan *hyperplane* yang berbeda maupun tidak tergantung dari data set yang dimiliki. Hal ini juga bisa dikarenakan SVM telah mengklasifikasi-kan data *training* secara tepat sehingga memperbesar maupun memperkecil nilai C tidak akan berpengaruh terhadap akurasi.

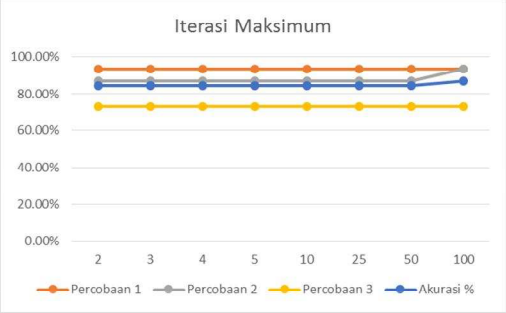
5.4 Pengujian dan Analisis Parameter Gamma



Gambar 9. Grafik Hasil Uji Parameter Gamma

Grafik menunjukkan bahwa pengujian terhadap nilai konstanta *gamma* memperoleh hasil akurasi rata – rata tertinggi pada nilai 0,001; 0,01; 0,1; dan 0,2 mencapai 86,67%. Dari grafik tersebut dapat dijelaskan bahwa konstanta *gamma* memiliki pengaruh terhadap tingkat akurasi. Semakin besar nilai *gamma*, maka akurasi akan semakin menurun karena nilai *gamma* yang tinggi menyebabkan *hyperplane* hanya memperhatikan data terdekat saja sedangkan data yang jauh akan diabaikan sehingga data *training* tidak terpetakan dengan baik (*underfitting*). Sedangkan semakin kecil nilai *gamma*, titik terjauh akan semakin diperhatikan sehingga mampu menyebabkan *overfitting*.

5.5 Pengujian dan Analisis Iterasi Maksimum



Gambar 10. Grafik Hasil Uji Iterasi Maksimum

Grafik menunjukkan bahwa akurasi terbaik terletak pada iterasi ke 100. Rata – rata akurasi yang diperoleh pada iterasi ke 100 mencapai 86,66 %. Dari grafik tersebut dijelaskan bahwa iterasi maksimum cenderung stabil dan semakin besar jumlah iterasi maka semakin besar kemungkinan untuk memperoleh nilai α_i yang konvergen akan semakin besar. Namun, iterasi maksimum tidak hanya dipengaruhi oleh jumlah maksimum iterasi melainkan juga jumlah ε (*epsilon*) di mana iterasi akan berhenti jika nilai ($|\delta\alpha| < \varepsilon$).

6. Kesimpulan dan Saran

6.1 Kesimpulan

Penggunaan metode ekstraksi fitur warna modus dan *mean* untuk klasifikasi dengan metode *Support Vector Machine* berhasil melakukan klasifikasi pada citra tanah ke dalam lima kelas yang telah ditentukan sebelumnya yaitu 5 Y 3/1, 10 YR 3/2, 10 YR 5/4, 2.5 YR 4/2, dan 7.5 YR 5/6. Tingkat keberhasilan deteksi warna tanah menggunakan metode *Support Vector Machine* ini berhasil memperoleh nilai 93,33% menggunakan fitur modus dan 93,33% menggunakan fitur *mean*. Parameter – parameter yang digunakan antara lain *lamda* = 0,5, *gamma* = 0,01, *complexity* = 1, *epsilon* = 0,001, dan iterasi maksimum = 3. Nilai parameter – parameter pada metode klasifikasi SVM yang diubah – ubah ternyata memberikan dampak terhadap perubahan nilai akurasinya. Dari pengujian pada parameter – parameter SVM yang telah dilakukan didapatkan nilai terbaik untuk parameter *lamda* adalah 0,1 – 0,5 dan 1 – 3, untuk parameter C tidak terdapat perubahan dengan rata – rata akurasi adalah 84,44%, nilai terbaik untuk parameter konstanta *gamma* adalah 0,001, 0,01, 0,1, dan 0,2, sedangkan nilai terbaik untuk jumlah iterasi adalah 100.

6.2 Saran

1. Diharapkan pada penelitian selanjutnya dapat mengklasifikasikan warna tanah ke dalam index warna kelas yang lebih bervariasi.
2. Diharapkan pada penelitian yang selanjutnya dilakukan penambahan terhadap variasi ekstraksi fitur untuk pendeteksian warna tanah tidak hanya modus dan *mean* RGB dari citra.
3. Diharapkan untuk penelitian selanjutnya dapat menggunakan kernel yang berbeda karena kernel yang lain memiliki karakter data yang berbeda.

Daftar Pustaka:

Sutanto,Rachman, (2005), "*Dasar - Dasar Ilmu Tanah*", Yogyakarta: Kanisius.

Jordan, Antonio, (2014), *Soil Color Never Lies* [online]. Available: http://blogs.egu.eu/divisions/sss/2014/03/30/soil-color-never-lies/.

Hardjowigeno, S, (1992), "*Ilmu Tanah*", Edisi Ketiga, Jakarta: PT. Mediyatama Sarana Perkasa.

Hardjowigeno, S, (1987), "*Ilmu Tanah*", Edisi Pertama, Jakarta: PT. Mediyatama Sarana Perkasa.

Madjid, A., (2007), "*Dasar-Dasar Ilmu Tanah*", Bahan Ajar Online Fakultas Pertanian Unsri.

Hermawati, Fajar Astuti, (2013), "*Pengolahan Citra Digital*", Yogyakarta: Penerbit ANDI.

Rizal, Achmad, (2014), *Pengolahan Citra* [online], Available: http://achmadrizal.staff.telkomuniversity.ac.id/ pengolahan-citra/.

Robbani, Ihwanudien Hasan, et al., (2016), "*Aplikasi Mobile Scotect : Aplikasi Deteksi Warna Tanah dengan Teknologi Citra Digital pada Android*", JTIIK, vol. 3, No.1, pp.19-26, April 2016.

Woods, Richard E., Gonzales, Rafael E., (2008), "*Digital Image Processing*", Third Edition, New Jersey: Pearson Education.

Kusuma, Yoke, (2016), *Implementasi Metode Support Vector Machine (SVM) Untuk Identifikasi Penyakit Daun Tanaman Kacang Tanah Menggunakan Gray-Level Coocurance Matrix (GLCM)*. Disertasi Sarjana Komputer pada FTIIK Universitas Brawijaya : tidak diterbitkan.

Sembiring, Krisantus, (2007), Penerapan Teknik Support Vector Machine untuk Pendeteksian Intrusi pada Jaringan [online]. Available : http://sutikno.blog.undip.ac.id/ files /2011 / 11 / tutorial – svm – bahasa – indonesia – oleh - krisantus.pdf.