

Implementasi Ekstraksi Ciri Histogram dan K-Nearest Neighbor untuk Klasifikasi Jenis Tanah di Kota Banjar, Jawa Barat

Rudiono¹, Donny Avianto²

^{1,2} Program Studi Teknik Informatika, Fakultas Teknologi Informasi dan Elektro
Universitas Teknologi Yogyakarta
Jl. Siliwangi (Ringroad Utara) Jombor Sleman, D.I. Yogyakarta 55285, Indonesia
Email: ¹rudi.ono@student.uty.ac.id, ²donny@staff.uty.ac.id

Masuk: 2 April 2019 ; Direvisi: 18 Oktober 2019; Diterima: 23 Oktober 2019

Abstract. Land plays an essential role in the availability of nutrients and water to support our life on earth. Soil quality can be observed based on its color and texture characteristics. By knowing the quality of the soil, the most suitable plants for planting can be determined. This study is conducted to examine the soil quality in Langensari. The most regions in Langensari are in the altitude of fewer than 25 meters above sea level that they are very potential for agriculture and plantation. The proposed system used in this research is a cross-sectional image of the ground as input. The image is then extracted using histogram feature extraction to obtain the intensity, standard deviation, skewness, energy, entropy and smoothness values. K-Nearest Neighbor is then used to classify the results. The proposed system was tested using 20 test images. Based on the experiment result, the system can classify soil types appropriately with accuracy reaching 60% when value of $K = 1$ and $K=3$.

Keywords: Soil Types Classification, Histogram Feature Extraction, K-Nearest Neighbor, Website.

Abstrak. Tanah memegang peranan penting dalam tersedianya unsur hara dan air bagi kehidupan makhluk hidup di bumi. Kualitas tanah dapat diketahui dari karakteristik warna dan teksturnya. Dengan mengetahui kualitas tanah, jenis tanaman yang paling tepat untuk ditanam dapat ditentukan. Penelitian ini mengenai kualitas tanah di Langensari. Sebagian besar wilayah Langensari dipilih karena memiliki ketinggian kurang dari 25 mdpl dimana sangat berpotensi sebagai daerah pertanian dan perkebunan. Sistem yang diusulkan menggunakan citra penampang tanah sebagai inputan. Citra kemudian diekstrak menggunakan ekstraksi ciri histogram untuk mendapatkan nilai intensitas, standar deviasi, skewness, energi, entropi, dan smoothness. Fitur yang dihasilkan kemudian diklasifikasikan menggunakan algoritma K-Nearest Neighbor. Sistem yang diusulkan diuji menggunakan 20 citra uji. Berdasarkan hasil pengujian, sistem mampu mengklasifikasikan jenis tanah secara tepat dengan akurasi mencapai 60% saat nilai $K = 1$ dan nilai $K=3$.

Kata Kunci: Klasifikasi Jenis Tanah, Ekstraksi Ciri Histogram, K-Nearest Neighbor, , Website.

1. Pendahuluan

Langensari merupakan salah satu kecamatan di Kota Banjar, Jawa Barat dengan topografi yang sangat beragam. Berdasarkan data yang dirilis oleh Badan Pusat Statistik pada tahun 2017, 36,95% dari total luas lahan yang dipergunakan untuk kegiatan pertanian dan perkebunan Kota Banjar terletak di Kecamatan Langensari [1]. Letak Kecamatan Langensari yang berada di ketinggian kurang dari 25 mdpl membuat daerah ini sangat cocok dijadikan sebagai daerah pertanian dan perkebunan. Wilayah Langensari sendiri terbagi ke dalam 4 dusun dengan jenis tanah yang berbeda-beda. Dusun yang pertama yaitu Dusun Kedungwaringin, memiliki jenis tanah humus dan tanah litosol. Dusun kedua yaitu Dusun Babakan, memiliki jenis tanah litosol. Sedangkan Dusun ketiga dan keempat yaitu Dusun Madura dan Dusun Langensari, masing-masing memiliki jenis tanah podsolik merah kuning dan aluvial.

Namun, minimnya pengetahuan masyarakat tentang jenis-jenis tanah mengakibatkan pemanfaatan tanah menjadi kurang maksimal. Hal ini dapat dilihat dari beberapa contoh seperti menanam padi ladang di tanah podsolik merah kuning, dimana hasil tanaman padi ladang tidak

tumbuh subur dan tanaman padi yang tidak sama tinggi. Jenis tanaman yang kurang sesuai dengan jenis tanah di wilayah tersebut selanjutnya bisa berdampak pada tidak maksimalnya hasil panen bahkan hingga ke gagal panen. Hal ini diperkuat dengan data yang dihimpun oleh Dinas Pertanian dan Ketahanan Pangan Kota Banjar pada kurun waktu 2013-2014 yang menunjukkan penurunan jumlah panen dari 50.749 ton di tahun 2013 menjadi 47.237 ton pada tahun 2014 [2]. Oleh karena itu, pengetahuan jenis tanah menjadi hal yang penting bagi para petani agar dapat menyesuaikan jenis tanaman dengan jenis tanah yang akan ditanami.

Salah satu cara untuk mengetahui jenis tanah adalah melalui warna tanah. Akan tetapi banyaknya variasi warna tanah memberikan tantangan tersendiri dalam menentukan jenis tanah. Acuan yang selama ini digunakan oleh para ahli untuk menentukan jenis tanah adalah mencocokkan warna tanah di suatu daerah dengan warna-warna tanah yang ada di dalam buku *Munsell Soil Color Chart* [3]. Cara ini tentu saja sangat tidak efektif terutama bagi golongan petani di daerah karena membutuhkan waktu lama dan ketelitian ekstra untuk menentukan jenis tanah.

Berdasarkan permasalahan di atas, penelitian kali ini akan mengusulkan suatu sistem yang dapat melakukan klasifikasi jenis-jenis tanah secara mudah dan cepat berdasarkan citra penampang tanahnya. Sistem yang diusulkan telah dibuat agar mampu mengenali jenis-jenis tanah yang ada di wilayah Langensari sehingga dapat memudahkan para petani yang berada di daerah tersebut dalam mengklasifikasikan jenis tanah. Sistem yang diusulkan memanfaatkan metode ekstraksi ciri histogram dalam tahap *preprocessing* citra tanah untuk nilai-nilai fitur dari data citra tanah yang akan diklasifikasi. Selanjutnya nilai-nilai fitur akan diklasifikasikan menggunakan algoritma *K-Nearest Neighbor* (KNN) dan perhitungan jarak *Euclidean* untuk menentukan jenis tanah dari citra yang diinputkan.

2. Tinjauan Pustaka

Metode KNN telah banyak digunakan untuk melakukan klasifikasi citra di berbagai bidang seperti: pengenalan penyakit jantung iskemik [4], pengenalan rambu-rambu lalu lintas [5], pengenalan jenis dan tingkat kematangan buah tomat [6], hingga ke pengenalan warna kulit wajah [7]. Banyaknya bidang penelitian yang menggunakan KNN untuk mengenali inputan berupa citra menunjukkan bahwa ide di balik metode ini cukup sederhana dan mudah untuk diimplementasikan [8]. Pada penelitian kali ini, metode KNN akan digunakan untuk mengenali jenis tanah sehingga petani dapat menentukan tanaman yang tepat untuk ditanam di tanah tersebut. Penelitian tentang pengenalan jenis tanah pernah dilakukan oleh beberapa peneliti di luar negeri seperti Dewangan dan Dedhe yang menggunakan beberapa metode pengolahan citra dan *Support Vector Machine* (SVM) untuk mengklasifikasi citra tanah ke dalam 7 kelas yaitu: *clay*, *clayey peat*, *clayey sand*, *humus clay*, *peat*, *sandy clay*, dan *silty clay* [9]. Penggunaan beberapa teknik pengolahan citra dan SVM juga pernah dilakukan oleh Barman dan Choudhury di dalam penelitiannya yang bertujuan untuk mengklasifikasi tekstur tanah. Tekstur tanah menjadi faktor utama yang perlu dipertimbangkan sebelum memulai suatu budidaya tanaman [10]. Penelitian Honawad et al. juga menguatkan bahwa pengenalan jenis tanah sangat penting dalam menentukan tanaman yang tepat untuk dibudidayakan di lahan tersebut. Pada penelitiannya, metode pengolahan citra *low mask*, *gabor filter*, dan *color quantization* digunakan untuk melakukan ekstraksi ciri dari citra tanah. Selanjutnya ciri tersebut akan di klasifikasi ke dalam salah satu dari 10 kelas menggunakan prosedur *similarity checking* [11].

Peneliti dari dalam negeri juga sudah pernah melakukan penelitian terkait pengenalan jenis tanah. Seperti penelitian yang dilakukan oleh Bimantoro tentang klasifikasi tanah humus berdasarkan citra RGB menggunakan metode *K-Nearest Neighbor*. Penelitian ini bertujuan menentukan antara tanah humus dan tanah non humus. Dari total 22 data uji dengan pembagian 67% data sampel dan 33% data uji didapatkan hasil akurasi dari sistem sebesar 100% [12]. Penelitian kedua dilakukan oleh Robbani, Trisnawati dan Noviyanti tentang deteksi warna tanah dengan menggunakan algoritma *K-Means* dengan *preprocessing* menggunakan Segmentasi *K-Means* dan median filter. Hasil pencocokan 276 data uji terhadap data latih, diketahui bahwa terdapat 26 data yang hasil klasifikasinya tidak sesuai, dan akurasi dari

program yang dihasilkan sebesar 90,58% [13]. Penelitian ketiga dilakukan oleh Indrawaty dan Tristiyono tentang pengenalan jenis tanah untuk lahan pertanian dengan menggunakan metode Euclidean Distance dan Analytical Hierarchy Process (AHP). Pada tahap preprocessing digunakan metode Gray Level Co-occurrence Matrices (GLCM) untuk mengetahui nilai rata-rata dari tanah yang telah diakuisisi. Berdasarkan hasil penelitian, sistem dapat melakukan klasifikasi dengan akurasi sebesar 80% dari lima sampel data uji [14].

2.1. Tanah

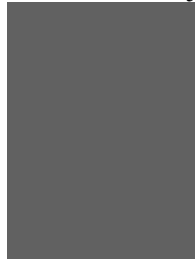
Tanah adalah campuran kompleks antara mineral, air, udara, bahan organik, dan sisa-sisa pembusukan organisme yang tak terhitung jumlahnya [15]. Campuran dari semua itu kemudian membentuk kulit bumi yang mampu mendukung kehidupan tanaman dan sangat penting bagi kehidupan di bumi. Tanah yang ada di bumi ini memiliki jenis yang berbeda-beda. Beberapa faktor yang mempengaruhi perbedaan jenis tanah antara lain adalah: bahan induk, iklim, waktu, dan keanekaragaman hayati termasuk aktivitas manusia [21, 24]. Perbedaan pada jenis tanah ini berdampak pula pada jenis tanaman yang mampu hidup di tanah tersebut. Oleh karena itu, klasifikasi jenis tanah sangat penting untuk meminimalkan salah tanam pada suatu daerah.

2.2. Jenis-Jenis Tanah

Jenis tanah yang ada di Indonesia bermacam-macam [16], antara lain :

1. Tanah Litosol

Tanah litosol merupakan jenis tanah yang terbentuk dari batuan beku yang berasal dari proses meletusnya gunung berapi dan juga sedimen keras dengan proses pelapukan kimia (dengan menggunakan bantuan organisme hidup) dan fisika (dengan bantuan sinar matahari dan hujan) yang belum sempurna. Warna tanah litosol ditunjukkan pada Gambar 1.

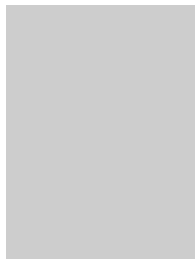


Gambar 1. Warna Tanah Litosol

Source: Munsell Color Chart Book

2. Tanah Aluvial

Tanah Aluvial merupakan tanah endapan, dibentuk dari lumpur dan pasir halus yang mengalami erosi tanah. Banyak terdapat di dataran rendah, di sekitar muara sungai, rawa-rawa, lembah-lembah, maupun di kanan kiri aliran sungai besar. Tanah ini banyak mengandung pasir dan liat, tidak banyak mengandung unsur-unsur zat hara. Warna tanah aluvial ditunjukkan pada Gambar 2.



Gambar 2. Warna Tanah Aluvial

Source: Munsell Color Chart Book

3. Tanah Podsolik Merah Kuning

Tanah Podsolik Merah Kuning adalah tanah yang mempunyai perkembangan profil, konsistensi teguh, bereaksi masam, dengan tingkat kejenuhan basa rendah. Podsolik merupakan segolongan tanah yang mengalami perkembangan profil dengan batas horizon yang jelas, berwarna merah hingga kuning dengan kedalaman satu hingga dua meter. Di wilayah tersebut tanah podsolik ini biasanya digunakan sebagai tanah untuk berkebun. Warna tanah podsolik merah kuning ditunjukkan pada Gambar 3.

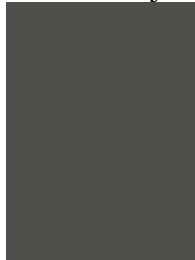


Gambar 3. Warna Tanah Podsolik Merah Kuning

Source: Munsell Color Chart Book

4. Tanah Humus

Tanah humus merupakan tanah yang terbentuk dari proses pelapukan dedaunan atau ranting- ranting pohon yang berjatuhan ke tanah. Proses terbentuknya tanah humus ini disebut dengan proses humifikasi. Warna tanah humus ditunjukkan pada Gambar 4.



Gambar 4. Warna Tanah Humus

Source: Munsell Color Chart Book

2.3. Citra

Citra adalah suatu gambaran atau kemiripan dari suatu objek. Akuisisi citra adalah proses pengubahan citra analog menjadi citra digital yang diambil dari lingkungan atau dunia nyata menggunakan beberapa alat berikut yaitu kamera digital, *webcam*, *smartphone*, scanner, dan lain-lain, agar bisa dilanjutkan ke tahap *preprocessing*. Sebelum masuk ke tahap *preprocessing* citra terlebih dahulu dilakukan beberapa proses yaitu *resize* dan *grayscale*. *Resize* adalah proses untuk memperbesar dan memperkecil ukuran citra dalam format piksel [22]. Sedangkan *grayscale* adalah proses mengubah citra RGB menjadi citra *grayscale* dengan cara membandingkan intensitas piksel pada satu saluran tertentu dengan intensitas total di setiap saluran warna [17].

2.4. Preprocessing

Preprocessing adalah langkah awal yang harus dilakukan pada citra, dengan melibatkan operasi tertentu dalam rangka memberikan dasar yang diperlukan untuk melakukan tugas selanjutnya [23]. Hal ini dilakukan karena hasil citra digital dari proses akuisisi biasanya memiliki beberapa masalah, misalnya terjadi noise atau adanya objek-objek pengganggu. Beberapa masalah tersebut disebabkan oleh kurang akuratnya sensor atau *transducer* yang digunakan saat proses akuisisi. Selain itu, tujuan *preprocessing* adalah untuk membuat citra digital agar sesuai dengan kebutuhan ekstraksi fiturnya.

2.5. Ekstraksi Ciri Histogram

Metode ekstraksi ciri histogram termasuk ke dalam kelompok metode statistik orde 1 karena berdasar pada karakteristik histogram citra [18]. Metode ini umumnya digunakan

mengekstraksi ciri tekstur dari suatu citra. Fitur pertama yang dihitung secara statis adalah rerata intensitas. Komponen fitur ini dihitung berdasarkan persamaan (1).

$$m = \sum_{i=0}^{L-1} i \cdot p(i) \quad (1)$$

Dalam hal ini, i adalah aras keabuan pada citra m dan $p(i)$ menyatakan probabilitas kemunculan dan L menyatakan nilai aras keabuan tertinggi.

Fitur kedua yang dihitung yaitu standar deviasi. Rumus perhitungannya ditunjukkan pada persamaan (2).

$$\sigma = \sqrt{\sum_{i=1}^{L-1} (i - m)^2 p(i)} \quad (2)$$

Dalam hal ini σ dinamakan standar deviasi, dimana i merupakan aras keabuan citra, L menyatakan nilai aras keabuan tertinggi, m merupakan rerata intensitas dan $p(i)$ merupakan fungsi peluang.

Fitur ketiga yang dihitung yaitu *skewness* merupakan ukuran ketidaksimetrisan terhadap rerata intensitas. Rumus perhitungannya ditunjukkan pada persamaan (3).

$$\text{skewness} = \sum_{i=1}^{L-1} (i - m)^3 p(i) \quad (3)$$

Dalam hal ini, i adalah aras keabuan pada citra, m merupakan rerata intensitas, $p(i)$ menyatakan probabilitas kemunculan dan L menyatakan nilai aras keabuan tertinggi.

Fitur keempat yang dihitung yaitu *descriptor* energi adalah yang menyatakan distribusi intensitas piksel terhadap jangkauan aras keabuan. Rumus perhitungannya ditunjukkan pada persamaan (4).

$$\text{energi} = \sum_{i=0}^{L-1} [p(i)]^2 \quad (4)$$

Dalam hal ini, i adalah aras keabuan pada citra, $p(i)$ menyatakan probabilitas kemunculan dan L menyatakan nilai aras keabuan tertinggi.

Fitur kelima yang dihitung yaitu entropi mengindikasikan kompleksitas citra. Rumus perhitungannya ditunjukkan pada persamaan (5).

$$\text{entropi} = \sum_{i=0}^{L-1} p(i) \log_2(p(i)) \quad (5)$$

Dalam hal ini, i adalah aras keabuan pada citra, $p(i)$ menyatakan probabilitas kemunculan dan L menyatakan nilai aras keabuan tertinggi.

Fitur keenam yang dihitung yaitu *smoothness*/properti kehalusan biasa disertakan untuk mengukur tingkat kehalusan atau kekerasan intensitas pada citra. Rumus perhitungannya ditunjukkan pada persamaan (6).

$$R = 1 - \frac{1}{1 + \sigma^2} \quad (6)$$

Dalam hal ini, R adalah tingkat kehalusan atau kekasaran pada citra dan σ^2 menyatakan nilai varians.

2.6. Klasifikasi Jenis Tanah

Secara umum klasifikasi dapat diartikan sebagai proses untuk menempatkan suatu objek ke dalam suatu kategori atau kelas yang sudah didefinisikan sebelumnya menggunakan model tertentu. Sedangkan klasifikasi jenis tanah adalah proses menentukan jenis tanah berdasarkan karakteristik tanah seperti warna dan teksturnya [19]. Klasifikasi tanah telah menjadi masalah penelitian yang menarik di bidang teknik [25]. Hal ini disebabkan karena hasil output dari klasifikasi jenis tanah ini mampu menghindarkan petani dari kerugian yang ditimbulkan dari kesalahan dalam memilih jenis tanaman untuk ditanam pada suatu daerah tertentu.

2.7. Algoritma *K-Nearest Neighbor*

Algoritma *K-Nearest Neighbor* (K-NN) adalah salah satu algoritma klasifikasi yang paling sering digunakan untuk memprediksi kelas dari catatan atau sampel dengan kelas yang tidak ditentukan berdasarkan kelas dari catatan tetangganya [20]. Pendekatan yang digunakan pada algoritma ini adalah perbandingan jarak kedekatan antara kasus baru dengan kasus lama berdasarkan kecocokan bobot dari sejumlah fitur yang ada yang memiliki kesamaan (*similarity*). Tujuan dari algoritma ini untuk mengklasifikasikan objek baru berdasarkan atribut dan *training sample*.

Ada banyak cara untuk mengukur jarak kedekatan antara data baru dengan data lama (*data training*), diantaranya menggunakan *euclidean distance* dan *manhattan distance* (*city block distance*), yang sering digunakan adalah *euclidean distance*, rumus perhitungan *euclidean distance* dihitung dengan menggunakan persamaan di bawah ini, yaitu :

$$\text{jarak} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2} \quad (7)$$

Dimana : a_1, a_2 dan $b : b_1, b_2$ mewakili nilai atribut dari dua record.

Menurut [20], langkah-langkah yang dilakukan pada algoritma *K-Nearest Neighbor* adalah sebagai berikut:

1. Menentukan parameter nilai K (jumlah tetangga terdekat).
2. Menghitung kuadrat jarak *euclidean* masing-masing objek terhadap data latih yang diberikan.
3. Mengurutkan jarak tersebut ke dalam kelompok yang mempunyai jarak *euclidean* terkecil.
4. Mengumpulkan kategori Y (Klasifikasi *K-Nearest Neighbor*) berdasarkan nilai K.
5. Dengan menggunakan kategori *K-Nearest Neighbor* yang paling mayoritas maka akan menghasilkan kelas data baru (prediksi).

Sedangkan untuk menghitung akurasi sistem pada penelitian ini digunakan persamaan (8) seperti di bawah ini :

$$\text{Akurasi} = \frac{\text{Jumlah Citra Dikenali}}{\text{Jumlah Citra Uji}} \times 100\% \quad (8)$$

Selain akurasi, persentase *error* juga akan dihitung menggunakan persamaan (9) berikut:

$$\text{Error} = \frac{\text{Jumlah Citra Tidak Dikenali}}{\text{Jumlah Citra Uji}} \times 100\% \quad (9)$$

3. Metodologi Penelitian

Metodologi yang digunakan dalam penelitian ini, yaitu:

3.1. Observasi

Dalam observasi penelitian ini, kegiatan pengamatan dilakukan di wilayah Langensari, Kota Banjar, Provinsi Jawa Barat. Observasi yang dilakukan mencakup tiga tahap yaitu tahap pengamatan, tahap pengambilan data tanah, dan tahap identifikasi citra.

1. Tahap Pengamatan

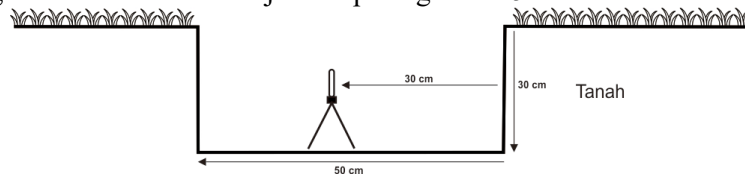
Tahap pengamatan dilakukan dengan cara mengamati dan melakukan survei berdasarkan sumber data. Adapun hasil dari pengamatan tanah berdasarkan daerahnya ditunjukkan pada tabel 1.

Tabel 1. Hasil Pengamatan Data Jenis Tanah

| No | Nama Daerah | Jenis Tanah |
|----|----------------|-------------------------------|
| 1 | Kedungwaringin | Tanah Humus dan Tanah Litosol |
| 2 | Babakan | Tanah Litosol |
| 3 | Madura | Tanah Podsolik Merah Kuning |
| 4 | Langensari | Tanah Aluvial |

2. Tahap Pengambilan Data Citra

Pada tahap pengambilan data citra tanah dilakukan dengan menggunakan kamera smartphone. Dalam proses pengambilan data citra tanah, dilakukan dengan cara membuat sebuah lubang dengan kedalaman 30 cm, panjang 50 cm, dan lebar 50 cm. kemudian kamera diletakan di lubang tersebut pada jarak 30 cm dari sisi tanah. Pengambilan gambar hanya dilakukan pada tanah yang kering dengan pengambilan data sebanyak 100 sampel data citra tanah. Ilustrasi pengambilan data citra ditunjukkan pada gambar 5.



Gambar 5. Ilustrasi Pengambilan Data Citra Tanah

3. Tahap Identifikasi Data Citra

Pada tahap identifikasi ini dilakukan dengan memilah gambar yang berkualitas baik dan memberi nama sesuai dengan jenis tanahnya. Berdasarkan tahap identifikasi maka didapatkan data citra tanah sebanyak 100 data. Data dikelompokkan menjadi 80 data citra untuk data latih dan 20 data citra untuk data uji. Kemudian dari data latih sebanyak 80 data dikategorikan ke dalam 4 jenis tanah dengan rincian 20 data untuk masing-masing jenis tanah, yaitu 20 data tanah litosol, 20 data tanah alluvial, 20 data tanah podsolik merah kuning dan 20 data tanah humus. Sedangkan data uji sebanyak 20 data dikategorikan ke dalam 4 jenis tanah dengan rincian, yaitu 5 data tanah litosol, 5 data tanah alluvial, 5 data tanah podsolik merah kuning dan 5 data tanah humus.

3.2. Studi Literatur

Kegiatan studi literatur dilakukan dengan mendalami tentang citra jenis tanah pada berbagai sumber seperti buku, dan mempelajari penelitian sebelumnya terkait dengan klasifikasi menggunakan algoritma K-Nearest Neighbor (KNN).

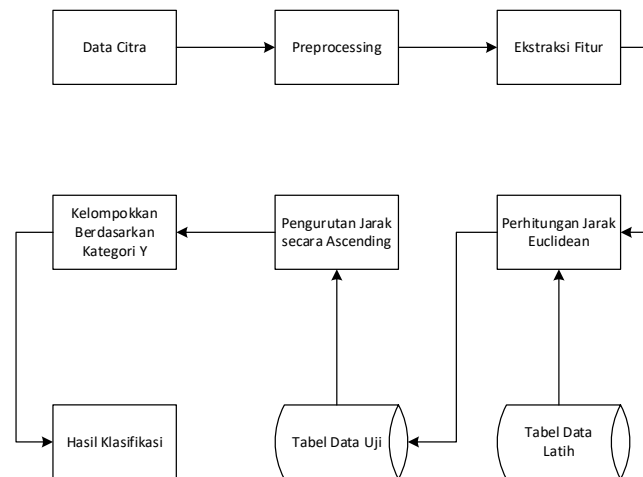
3.3. Perancangan Sistem

Pada tahap ini dilakukan tahap untuk menspesifikasikan bagaimana sistem dapat memenuhi kebutuhan informasi dan kebutuhan pengguna. Sistem ini akan memerlukan beberapa tahap desain seperti desain masukan, desain basis data, desain proses, desain keluaran, dan desain antarmuka. Selain itu pada desain sistem akan diberikan gambaran secara detail tentang Diagram Alir Data (DAD) dan Entity Relationship Diagram (ERD).

4. Hasil dan Diskusi

4.1. Gambaran Umum Sistem

Sistem klasifikasi jenis tanah merupakan sebuah sistem berbasis website yang dibangun untuk mengembangkan sistem yang dapat melakukan proses identifikasi dan klasifikasi jenis tanah berdasarkan citra tanah yang dimasukkan ke dalam sistem. Secara umum jalannya sistem klasifikasi jenis tanah dapat dilihat pada Gambar 2.



Gambar 6. Alur Kerja Sistem Klasifikasi Jenis Tanah

Berdasarkan gambar 2, cara kerja sistem klasifikasi jenis tanah dimulai dengan input data citra tanah yang telah diambil pada tahap akuisisi citra. Kemudian masuk ke tahap *preprocessing* sebelum masuk ke ekstraksi fitur. Pada tahap *preprocessing*, data citra tanah terlebih dahulu dilakukan *resize* dengan ukuran 416 x 312 piksel. Sebelum masuk ekstraksi fitur, citra akan diubah terlebih dahulu dari format RGB menjadi format *grayscale*. Selanjutnya, proses ekstraksi ciri akan diterapkan pada citra *grayscale* yang baru dihasilkan. Proses ekstraksi dimulai dengan menghitung fitur pertama yaitu rerata intensitas menggunakan Persamaan (1). Setelah Rerata intensitas (\bar{s}) didapatkan, nilai tersebut akan digunakan untuk mencari 2 fitur selanjutnya yaitu standar deviasi (σ) dan *skewness* menggunakan Persamaan (2) dan Persamaan (3). Langkah berikutnya adalah mencari nilai untuk 3 fitur lainnya yaitu energi, entropi, dan *smoothness* (R) masing-masing menggunakan Persamaan (4), Persamaan (5), dan Persamaan (6).

Setelah mendapatkan keenam nilai fitur dari citra, tahap selanjutnya adalah menghitung jarak antara data nilai fitur pada citra yang diinputkan dengan setiap nilai fitur pada data latih. Adapun penghitungan jarak pada penelitian ini menggunakan rumus *Euclidean distance* seperti pada Persamaan (7). Setelah seluruh jarak dari nilai fitur citra input ke setiap data latih ditemukan, nilai jarak ini akan diurutkan secara *ascending* (dari jarak terdekat ke jarak terjauh). Jarak yang telah diurutkan kemudian di kelompokkan berdasarkan kategori Y, yang mana kategori Y didapatkan dari nilai K yang ditetapkan pada saat perhitungan *K-Nearest Neighbor*. Nilai mayoritas yang terdapat pada kategori Y tersebut merupakan hasil dari klasifikasinya.

4.2. Proses Pelatihan

Pada tahap pelatihan, yang dilakukan adalah proses ekstraksi fitur citra data latih menggunakan metode ekstraksi ciri histogram. Tahap ini berfungsi untuk mendapatkan nilai fitur dari citra. Adapun nilai fitur yang di ekstraksi yaitu intensitas, standar deviasi, *skewness*, energi, entropi, dan *smoothness*.

Pada proses pelatihan digunakan 80 data yang terbagi ke dalam 4 jenis tanah yaitu tanah litosol, tanah aluvial, tanah podsolik merah kuning dan tanah humus. Dimana masing-masing jenis tanah tersebut memiliki 20 data. Proses pelatihan dilakukan untuk mendapatkan nilai K yang paling optimal, nilai K yang digunakan pada proses pelatihan yaitu 1,3,5,7. Proses pelatihan didapatkan akurasi tertinggi pada nilai K = 1 dan nilai K = 3 dengan tingkat akurasi sebesar 60%. Diberikan sebuah contoh akurasi yang didapatkan dari proses pelatihan dengan nilai K=1 dan nilai K =3 yang ditunjukkan pada tabel 2.

Tabel 2. Data Pelatihan

| No | Nama Tanah | Jumlah Data | Nilai K | Akurasi per jenis | Akurasi per nilai K |
|----|-----------------------|-------------|---------|-------------------|---------------------|
| 1 | Podsolik Merah Kuning | 5 | 1 | 80 | 60 |
| 2 | Aluvial | 5 | 1 | 60 | |
| 3 | Litosol | 5 | 1 | 80 | |
| 4 | Humus | 5 | 1 | 20 | |
| 5 | Podsolik Merah Kuning | 5 | 3 | 80 | 60 |
| 6 | Aluvial | 5 | 3 | 40 | |
| 7 | Litosol | 5 | 3 | 80 | |
| 8 | Humus | 5 | 3 | 40 | |

Berdasarkan tabel 2, proses pelatihan dilakukan pada empat jenis tanah dengan total 20 data untuk satu nilai K sehingga total data yang digunakan pada empat nilai K adalah 80 data. Pada tabel 2, ditunjukkan akurasi tertinggi terdapat pada tanah jenis podsolik merah kuning dan litosol dengan akurasi sebesar 80% sedangkan akurasi tertinggi untuk nilai K yaitu pada nilai K = 1 dan nilai K = 3 dengan akurasi sebesar 60%.

4.3. Proses Pengujian

Pada tahap proses pengujian dilakukan untuk mendapatkan nilai K yang paling tepat untuk proses klasifikasi. Dimana proses pengujian dilakukan dengan masing-masing nilai K dengan jumlah data untuk tiap nilai K sebanyak 20 data untuk mendapatkan akurasi yang paling tinggi pada nilai K, sehingga nilai K yang memiliki nilai akurasi tertinggi tersebut yang akan digunakan untuk proses klasifikasi. Nilai K yang digunakan yaitu 1,3,5 dan 7. Proses Pengujian pada masing-masing nilai K dapat dilihat pada tabel 3.

Tabel 3. Proses Pengujian

| No | Nama Tanah | Jumlah Data | Nilai K | Akurasi per jenis | Akurasi per nilai K |
|----|-----------------------|-------------|---------|-------------------|---------------------|
| 1 | Podsolik Merah Kuning | 5 | 1 | 80 | 60 |
| 2 | Aluvial | 5 | 1 | 60 | |
| 3 | Litosol | 5 | 1 | 80 | |
| 4 | Humus | 5 | 1 | 20 | |
| 5 | Podsolik Merah Kuning | 5 | 3 | 80 | 60 |
| 6 | Aluvial | 5 | 3 | 40 | |
| 7 | Litosol | 5 | 3 | 80 | |
| 8 | Humus | 5 | 3 | 40 | |
| 9 | Podsolik Merah Kuning | 5 | 5 | 60 | 45 |
| 10 | Aluvial | 5 | 5 | 40 | |
| 11 | Litosol | 5 | 5 | 60 | |
| 12 | Humus | 5 | 5 | 20 | |
| 13 | Podsolik Merah Kuning | 5 | 7 | 60 | 55 |
| 14 | Aluvial | 5 | 7 | 40 | |
| 15 | Litosol | 5 | 7 | 80 | |
| 16 | Humus | 5 | 7 | 40 | |

Berdasarkan tabel 3, pengujian dilakukan pada masing-masing nilai K dengan jumlah data pada masing-masing nilai K adalah 20 data. Dari hasil pengujian didapatkan bahwa akurasi tertinggi terdapat pada nilai K = 1 dan nilai K = 3 yaitu sebesar 60%.

Langkah-langkah proses pengujian yang dilakukan adalah sebagai berikut:

1. Langkah pertama adalah menentukan nilai parameter K. Dalam proses pengujian ini ditentukan nilai K yaitu 1,3, 5 dan 7.
2. Langkah kedua adalah mengunggah file citra yang digunakan untuk proses pengujian. File citra yang digunakan adalah file citra tanah hasil akuisisi citra yang kemudian di kelompokkan sesuai dengan jenis tanahnya.

3. Langkah ketiga adalah ekstraksi fitur. Dari langkah ini didapatkan nilai fitur dari citra tanah yang diunggah.
4. Langkah keempat adalah proses perhitungan jarak Euclidean menggunakan rumus pada persamaan (7).
5. Langkah kelima adalah mengurutkan data jarak secara ascending (dari jarak terkecil ke jarak terbesar).
6. Langkah keenam yaitu menentukan kategori Y sesuai dengan nilai K yang dipilih pada tahap pertama.

Dari hasil perhitungan langkah-langkah maka akan didapatkan hasil kesesuaian antara jenis tanah dengan hasil klasifikasinya.

Diberikan contoh perhitungan algoritma KNN dengan nilai $K = 1$ dan data citra tanah podsolik merah kuning yang memiliki nilai fitur yaitu intensitas = 118,873, standar deviasi = 36,6833, skewness = 221,214, energi = 0,0095, entropi = 4,9359, smoothness = 0,0203.

Kemudian diberikan data citra tanah baru yang belum diketahui jenis tanahnya dengan nilai fitur yaitu intensitas = 119,221, standar deviasi = 40,6395, skewness = 226,011, energi = 0,0081, entropi = 5,0352, smoothness = 0,0248.

Dari kedua data nilai fitur citra tersebut kemudian akan dihitung jarak Euclidean dengan menggunakan rumus pada persamaan (7). Contoh perhitungan jarak Euclidean dengan data uji yang pertama adalah sebagai berikut:

$$\begin{aligned} \text{jarak} &= \sqrt{(119.221 - 118.873)^2 + (40.6395 - 36.6833)^2 + (226.011 - 221.214)^2 +} \\ &\quad (0.0081 - 0.0095)^2 + (5.0352 - 4.9359)^2 + (0.0248 - 0.0203)^2} \\ \text{jarak} &= 6,2280 \end{aligned}$$

Dari hasil perhitungan data ke-1 sampai data ke-20 diperoleh jarak terpendek sebesar 6,2280. Dikarenakan nilai K yang ditentukan 1 maka nilai jarak terpendeklah yang diambil sebagai nilai terbaik. Dari hasil perhitungan data ke-1 sampai data ke-20 diketahui bahwa perhitungan data ke-1 mendapatkan jarak terpendek dan data ke-1 pada proses pengujian adalah tanah jenis podsolik merah kuning, sehingga citra tanah tersebut adalah tanah podsolik merah kuning.

4.4. Implementasi Sistem

1. Akuisisi Citra

Akuisisi citra dilakukan untuk mengubah data analog menjadi data digital dengan cara mengambil data langsung dari dunia nyata. Pengambilan data citra ini dilakukan dengan menggunakan perangkat kamera smartphone. Jumlah data citra yang diambil adalah sebanyak 100 data, yang terbagi ke dalam data pelatihan yaitu sebanyak 80 data dan untuk proses prediksi sebanyak 20 data. Data jenis tanah yang diambil ada 4 jenis, yaitu tanah humus, tanah alluvial, tanah podsolik merah kuning, dan tanah litosol. Hasil akuisisi citra tanah ditunjukkan pada gambar 6.

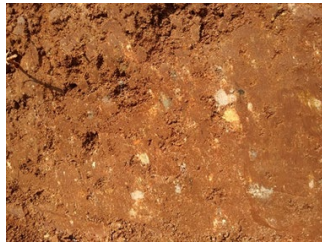


Gambar 7. Hasil Akuisisi Citra Tanah

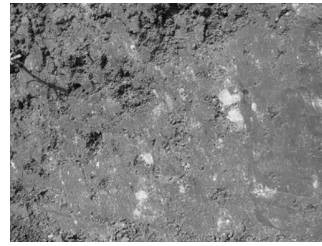
2. Preprocessing

Pada tahap *preprocessing* dilakukan dua tahap yaitu yang pertama adalah *resize* file citra menjadi ukuran 416 x 312 piksel. Hal ini dilakukan agar pada proses perhitungan nilai pada tiap piksel tidak terlalu lama karena ukuran piksel yang terlalu besar. Tahap kedua yaitu tahap

mengubah citra RGB menjadi citra *grayscale*. Hal ini dilakukan karena citra yang digunakan pada metode ekstraksi ciri histogram adalah citra *grayscale*. Citra RGB dan citra *grayscale* ditunjukkan pada gambar 7 dan gambar 8.



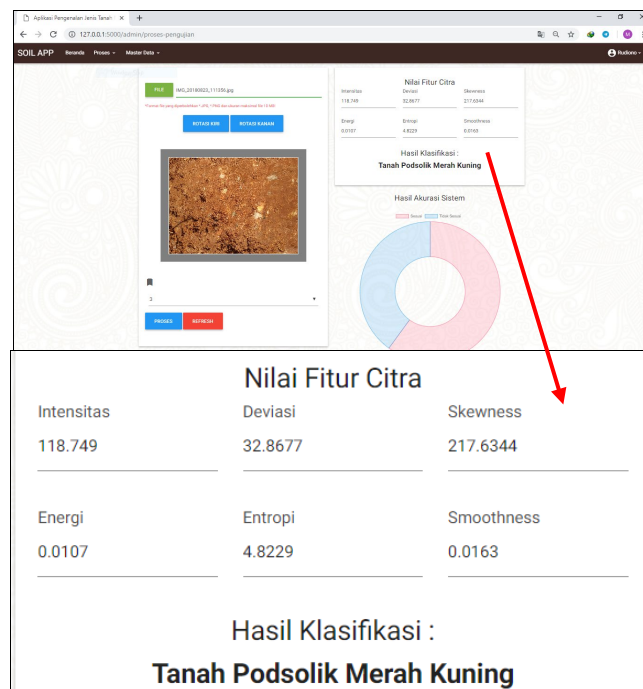
Gambar 8. Citra RGB



Gambar 9. Citra Grayscale

3. Ekstraksi Fitur dan Klasifikasi

Pada tahap ini citra digital yang telah melewati proses *preprocessing* diambil ciri objeknya yang bisa menjadi bahan pembeda dari objek-objek lainnya. Ekstraksi fitur merupakan informasi yang membedakan objek satu dengan yang lainnya yang mana kemudian digunakan sebagai parameter atau nilai masukan untuk membedakan objek pada tahap klasifikasi. Fitur atau ciri yang digunakan dalam penelitian ini adalah metode rerata intensitas, standar deviasi, *skewness*, energi, entropi dan *smoothness*. Pada tahap klasifikasi ditetapkan nilai $K = 3$. Nilai ekstraksi fitur citra dan hasil klasifikasi ditunjukkan pada gambar 10.



Gambar 10. Nilai Fitur Citra dan Hasil Klasifikasi

Berdasarkan gambar 10, nilai fitur citra telah didapatkan beserta hasil klasifikasi dari citra tanah yang dimasukkan ke dalam sistem. Hasil akurasi sistem klasifikasi jenis tanah sebesar 60% yang ditunjukkan oleh grafik berwarna merah dan *persentase error* sistem didapatkan sebesar 40% yang ditunjukkan oleh grafik berwarna biru. Hasil akurasi dan *persentase error* didapatkan dari hasil pengujian dengan jumlah 20 data dengan nilai $K = 3$. Hasil pengujian dengan nilai $K = 3$ ditunjukkan pada tabel 4.

Tabel 4. Hasil Pengujian dengan Nilai K = 3

| Citra Uji | Hasil Klasifikasi | Keterangan |
|-------------|-----------------------------|--------------|
| Citra ke-1 | Tanah Podsolik Merah Kuning | Sesuai |
| Citra ke-2 | Tanah Podsolik Merah Kuning | Sesuai |
| Citra ke-3 | Tanah Podsolik Merah Kuning | Sesuai |
| Citra ke-4 | Tanah Humus | Tidak Sesuai |
| Citra ke-5 | Tanah Podsolik Merah Kuning | Sesuai |
| Citra ke-6 | Tanah Litosol | Tidak Sesuai |
| Citra ke-7 | Tanah Litosol | Tidak Sesuai |
| Citra ke-8 | Tanah Humus | Tidak Sesuai |
| Citra ke-9 | Tanah Aluvial | Sesuai |
| Citra ke-10 | Tanah Aluvial | Sesuai |
| Citra ke-11 | Tanah Humus | Sesuai |
| Citra ke-12 | Tanah Humus | Sesuai |
| Citra ke-13 | Tanah Humus | Sesuai |
| Citra ke-14 | Tanah Humus | Sesuai |
| Citra ke-15 | Tanah Aluvial | Tidak Sesuai |
| Citra ke-16 | Tanah Litosol | Sesuai |
| Citra ke-17 | Tanah Aluvial | Tidak Sesuai |
| Citra ke-18 | Tanah Aluvial | Tidak Sesuai |
| Citra ke-19 | Tanah Humus | Tidak Sesuai |
| Citra ke-20 | Tanah Litosol | Sesuai |

Berdasarkan hasil pengujian pada tabel 2, didapatkan 12 data sesuai hasil klasifikasi dan 8 data tidak sesuai dengan hasil klasifikasi sehingga didapatkan persentase *error* pada proses pengujian sebesar 40%. Tingat persentase *error* tertinggi terjadi pada hasil pengujian tanah aluvial dan tanah litosol. Hal ini terjadi dikarenakan citra tanah berjenis aluvial dan litosol memiliki kemiripan yang tinggi dari segi warna. Kondisi ini menyebabkan nilai fitur yang dihasilkan kedua jenis tanah tersebut memiliki nilai yang hampir sama.

5. Kesimpulan dan Saran

Berdasarkan hasil penelitian yang telah dilakukan maka dapat disimpulkan bahwa sistem klasifikasi jenis tanah berbasis *website* menggunakan ekstraksi ciri histogram dan algoritma *K-Nearest Neighbor* (KNN) telah berhasil dibuat dan dapat menampilkan hasil klasifikasi jenis tanah. Selain itu, sistem dapat menampilkan nilai fitur dari citra yang diklasifikasi. Nilai fitur citra yang diekstraksi menggunakan ekstraksi ciri histogram yaitu intensitas, standar deviasi, *skewness*, energi, entropi, dan *smoothness*. Berdasarkan hasil pengujian, sistem memiliki nilai akurasi tertinggi sebesar 60% yaitu pada nilai $K = 1$ dan nilai $K=3$ dari total data uji sebanyak 20 data.

6. Ucapan Terima Kasih

Terima kasih penulis ucapkan kepada pihak Dinas Pertanian dan Ketahanan Pangan Kota Banjar yang telah membantu memberikan data terkait jenis-jenis tanah di Kota Banjar.

Referensi

- [1] E. A. Gandirum, Kota Banjar dalam Angka Tahun 2015. Banjar, 2015.
- [2] N. T. Basuki, Kota Banjar Dalam Angka 2018. Banjar: BPS Kota Banjar, 2018.
- [3] A. H. Munsell, Munsell Soil Book of Color. Michigan USA: X-Rite Incorporated, 2009.
- [4] Y. Udovychenko, A. Popov, and I. Chaikovsky, "Ischemic heart disease recognition by k-NN classification of current density distribution maps," in International Conference on Electronics and Nanotechnology (ELNANO), 2015.
- [5] Y. Han, K. Virupakshappa, and E. Oruklu, "Robust traffic sign recognition with feature extraction and k-NN classification methods," in International Conference on Electro/Information Technology (EIT), 2015.

- [6] W. D. N. Pacheco and F. R. J. López, "Tomato classification according to organoleptic maturity (coloration) using machine learning algorithms K-NN, MLP, and K-Means Clustering," in XXII Symposium on Image, Signal Processing and Artificial Vision (STSIVA), 2019.
- [7] C. Eyupoglu, "Implementation of color face recognition using PCA and k-NN classifier," in IEEE NW Russia Young Researchers in Electrical and Electronic Engineering Conference (EIconRusNW), 2016.
- [8] E. C. Ozan, E. Riabchenko, S. Kiranyaz, and M. Gabbouj, "A vector quantization based k-NN approach for large-scale image classification," in A vector quantization based k-NN approach for large-scale image classification, 2016.
- [9] P. Dewangan and V. Dedhe, "Soil Classification Using Image Processing and Modified SVM Classifier," *Int. J. Trend Sci. Res. Dev.*, vol. II, no. 6, pp. 504–507, 2018.
- [10] U. Barman and R. D. Choudhury, "Soil texture classification using multi class support vector machine," *Inf. Process. Agric.*, 2019.
- [11] S. K. Honawad, S. S. Chinchali, K. Pawar, and P. Deshpande, "Soil Classification and Suitable Crop Prediction," in National Conference On Advances In Computational Biology, Communication, And Data Analytics, 2017, pp. 25–29.
- [12] D. A. Bimantoro, "Klasifikasi Tanah Humus Berdasarkan Citra RGB Menggunakan Metode K-Nearest Neighbor (KNN)," *J. Sains dan Teknol.*, vol. IV, no. 7, pp. 12–27, 2016.
- [13] I. H. Robbani, E. Trisnawati, R. Noviyanti, A. Rivaldi, and F. Utaminingrum, "Aplikasi Mobile Scotect: Aplikasi Deteksi Warna Tanah Dengan Teknologi Citra Digital Pada Android," *J. Teknol. Inf. dan Ilmu Komput.*, vol. III, no. 1, pp. 19–26, 2016.
- [14] Y. I. N and A. Tristiyono, "Aplikasi Pengenalan Jenis Tanah Untuk Lahan Pertanian Dengan Menggunakan Metode Euclidean Distance," *MIND J.*, vol. I, no. 1, pp. 26–35, 2016.
- [15] V. M. Dolas and P. U. Joshi, "A Novel Approach for Classification of Soil and Crop Prediction," *Int. J. Comput. Sci. Mob. Comput.*, vol. 7, no. 3, pp. 20–24, 2018.
- [16] S. Hardjowigeno, *Ilmu Tanah, Edisi Ketiga*. Jakarta, Indonesia: PT. Mediyatama Sarana Perkasa, 2013.
- [17] K. Padmavathi and K. Thangadurai, "Implementation of RGB and Grayscale Images in Plant Leaves Disease Detection – Comparative Study," *Indian J. Sci. Technol.*, vol. 9, no. February, pp. 4–9, 2016.
- [18] F. Muwardi and A. Fadlil, "PENGOLAHAN CITRA DAN PENGKLASIFIKASI JARAK," *J. Ilmu Tek. Elektro Komput. dan Inform.*, vol. 3, no. 2, pp. 124–131, 2017.
- [19] M. Kaur and R. K. Randhawa, "Texture Based Classification of Indian Soils Using Local Binary Pattern and Artificial Neural Networks," *Int. J. Electron. Comput. Sci. Eng.*, vol. 4, no. 3, pp. 228–237, 1956.
- [20] M. R. Alghifari and A. P. Wibowo, "Penerapan Metode K-Nearest Neighbor Untuk Klasifikasi Kinerja Satpam Berbasis Web," *J. Teknol. Manaj. Inform.*, vol. V, no. 1, 2019.
- [21] S. A. Siddiqui and N. Fatima, "Indian Soils : Identification and Classification," *Earth Sci. India*, vol. 10, no. III, pp. 1–14, 2017.
- [22] S. Perumal and T. Velmurugan, "Preprocessing by Contrast Enhancement Techniques for Medical Images," *Int. J. Pure Appl. Math.*, vol. 118, no. 18, pp. 3681–3688, 2018.

- [23] H. R. Thaker and C. K. Kumbharana, "Study and Analysis of various preprocessing approaches to enhance Offline Handwritten Gujarati Numerals for feature extraction," *Int. J. Sci. Res. Publ.*, vol. 4, no. 7, pp. 1–5, 2014.
- [24] H. Obara, Y. Maejima, K. Kohyama, T. Ohkura, and Y. Takata, "Outline of the Comprehensive Soil Classification System of Japan – First Approximation," *Japan Agric. Res. Q.*, vol. 49, no. 3, pp. 217–226, 2015.
- [25] Chandan and R. Thakur, "Recent Trends Of Machine Learning In Soil Classification : A Review," *Int. J. Comput. Eng. Res.*, vol. VIII, no. 9, pp. 25–32, 2018.