

Perbandingan Hasil Klasifikasi Tutupan Lahan Menggunakan Metode Artificial Neural Network (ANN), Support Vector Machine (SVM), dan Random Forest (RF) dengan Bahasa Pemrograman R

Comparison of Land Cover Classification Results Using Artificial Neural Network (ANN), Support Vector Machine (SVM), And Random Forest (RF) Methods with R Programming Language

Klarissa Ardilia Putri, Hepi Hapsari Handayani*

Departemen Teknik Geomatika, FTSPK-ITS, Kampus ITS Sukolilo, Surabaya, 60111, Indonesia

*Korespondensi penulis: hapsari@geodesy.its.ac.id

Diterima: 30072022; Diperbaiki: 23042024; Disetujui: 18072024; Dipublikasi: 18072024

Abstrak: Teknologi penginderaan jauh telah mengalami perkembangan yang pesat terutama dalam metode klasifikasi citra. Klasifikasi citra merupakan kegiatan yang dilakukan untuk mendeteksi dan mengidentifikasi objek di permukaan bumi pada citra satelit. Saat ini telah dikembangkan berbagai metode klasifikasi citra satelit, salah satunya adalah metode klasifikasi berbasis parameter, seperti MLC (Maximum Likelihood). Namun, metode ini tidak dapat dilakukan di lingkungan dengan fitur objek yang kompleks (seperti di daerah perkotaan atau daerah yang dibangun dengan padat). Hal ini dikarenakan metode klasifikasi berdasarkan parameter yang dibutuhkan dari dataset tidak berdistribusi normal. Oleh karena itu, dikembangkan metode klasifikasi berbasis non-parametrik yang tidak bergantung pada sifat sebaran data, sehingga tidak diperlukan parameter statistik untuk memisahkan banyak kelas pada citra dan digunakan untuk menganalisis citra satelit dengan lahan padat dan kompleks. fitur penutup. Melalui penelitian ini dilakukan analisis komparatif klasifikasi tutupan lahan menggunakan metode klasifikasi berbasis non parametrik dengan pendekatan Machine Learning (*Artificial Neural Network, Support Vector Machine, dan Random Forest*) untuk wilayah Kota Surabaya menggunakan Landsat-8 data citra satelit, untuk mengetahui hasil klasifikasi kinerja yang dihasilkan dari ketiga metode tersebut. Klasifikasi tutupan lahan yang diterapkan pada penelitian ini terdiri dari empat kelas yaitu, badan air, lahan terbuka, lahan terbangun, dan vegetasi. Adapun komposisi training point yang digunakan adalah 80:20, dimana 80% titik sebagai titik sample dan 20% sebagai titik validasi. Dan jumlah seluruh training point adalah 237 titik. Hasil klasifikasi tutupan lahan pada penelitian ini kemudian dilakukan uji akurasi secara kualitatif dan kuantitatif. Berdasarkan uji akurasi secara kuantitatif metode Random Forest menunjukkan hasil yang paling baik dengan nilai overall accuracy 93,33% dan kappa accuracy sebesar 91,07%.

Copyright © 2024 Geoid. All rights reserved.

Abstract: Remote sensing technology has experienced rapid development, especially in the field of image classification method. Image classification is an activity carried out to detect and identify objects on the earth's surface in satellite images. Currently, various satellite image classification methods have been developed, one of which is a parameter-based classification method, such as MLC (Maximum Likelihood). However, this method cannot be performed in environments with complex object features (such as in urban areas or densely built areas). This is because the classification method based on the required parameters of the dataset is not normally distributed. Therefore, a non-parametric-based classification method has been developed that is independent of the nature of the data distribution, so that no statistical parameters are needed to separate many classes in the image and are used to analyse satellite images with dense and complex land cover features. Through this research, a comparative analysis of land cover classification was carried out using a non-parametric-based classification method with a Machine Learning approach (*Artificial Neural Network, Support Vector Machine, and Random Forest*) for the Surabaya City area using Landsat-8 satellite imagery data, to find performance classification results resulting from the third method. The land cover classification applied in this study consisted of four classes, there are water, bare land, urban land, and vegetation. The composition of the training points used is 80:20, which 80% points as sample points, and 20% as validation points. And the total number of training points is 237 points. The land cover classification results in this study were tested for accuracy quantitatively and qualitatively. Based on the quantitative accuracy test, the Random Forest method showed the best results with an overall accuracy value of 93.33% and a kappa accuracy of 91.07%.

Kata kunci: *Machine Learning, Artificial Neural Network, Random Forest, Support Vector Machine, Tutupan Lahan*

Cara untuk sitasi: Putri, K.A., & Handayani, H.H.. (2024). Analisis Hasil Klasifikasi Tutupan Lahan Menggunakan Metode Artificial Neural Network (ANN), Support Vector Machine (SVM), dan Random Forest (RF) dengan Bahasa Pemrograman R. *Geoid*, 19(2), 349 – 360.

Pendahuluan

Penginderaan jauh (*remote sensing*) merupakan ilmu yang mengacu pada identifikasi fitur permukaan bumi dan sifat geo-biofisikanya dengan menggunakan radiasi elektromagnetik sebagai media interaksinya. Data permukaan bumi yang terekam oleh sensor penginderaan jauh terdiri dari panjang gelombang yang berbeda-beda, sehingga harus dilakukan koreksi secara radiometrik dan geometrik sebelum data informasi spektral dapat diekstrak (Havyas dkk., 2015). Saat ini teknologi penginderaan jauh telah mengalami perkembangan yang pesat, yang diiringi dengan semakin luasnya lingkup aplikasi teknologi penginderaan jauh. Adapun tujuan dikembangkannya penginderaan jauh adalah untuk merekam data yang ada di permukaan bumi, yang selanjutnya digunakan untuk inventarisasi dan evaluasi pemanfaatan kekayaan alam yang tersimpan di bumi. Penelitian dengan menggunakan data citra satelit penginderaan jauh banyak dikembangkan dengan metode klasifikasi citra. Hal ini dikarenakan hasil klasifikasi citra dapat digunakan sebagai dasar untuk melakukan interpretasi, analisis, serta pemodelan data spasial yang dapat digunakan untuk membantu dalam penentuan berbagai kebijakan lingkungan dan sosial ekonomi (Jensen dkk., 1999). Teknologi penginderaan jauh dapat dijadikan sebagai salah satu alternatif dalam mendukung pengembangan sistem pemantauan perubahan tutupan lahan saat ini (Suryo, 2013). Prediksi perubahan tutupan lahan dapat dilakukan dengan menggunakan analisis berbasis spasial menggunakan data citra penginderaan jauh (Nugroho dan Handayani, 2021). Selain itu, dalam melakukan perencanaan tata ruang kota, pemerintah perlu mengetahui penggunaan area atau wilayah kota, sehingga perlu dibuat peta tutupan lahan yang memiliki peranan penting untuk mengidentifikasi bagian tutupan lahan perkotaan sehingga dapat dilakukan perencanaan tata ruang kota sesuai dengan fungsi lahan yang ada (Natural Resources, 2015)

Klasifikasi citra merupakan suatu kegiatan yang dilakukan untuk melakukan deteksi dan identifikasi terhadap obyek-obyek yang ada di permukaan bumi pada citra satelit. Dengan mengenali obyek-obyek tersebut melalui unsur-unsur utama spektral dan spasial serta kondisi temporalnya. Saat ini telah berkembang berbagai metode klasifikasi citra satelit, salah satunya adalah metode klasifikasi berbasis parameter (*Parametric Classification*), seperti MLC (*Maximum Likelihood*). Namun, pada metode ini tidak dapat dilakukan lingkungan dengan fitur objek yang kompleks (seperti pada area perkotaan atau daerah yang padat bangunan). Dikarenakan, pada metode klasifikasi berbasis parameter dibutuhkan dataset yang tidak terdistribusi normal. Oleh karena itu, mulai dikembangkan metode klasifikasi berbasis non parametrik (*Non-parametric Classifier*) yang bersifat independen terhadap sifat distribusi data, sehingga pada metode ini tidak diperlukan parameter statistik untuk memisahkan banyak kelas pada citra, sehingga metode ini sangat sesuai digunakan untuk menganalisis citra satelit dengan fitur tutupan lahan yang padat dan kompleks. Metode klasifikasi berbasis non parametrik yang sering digunakan adalah metode *Machine Learning Decision Trees*, *Random Forest*, *Artificial Neural Networks* dan *Support Vector Machines* (Havyas dkk., 2015).

Melalui penelitian ini, dilakukan analisis perbandingan klasifikasi tutupan lahan menggunakan pendekatan *Machine Learning* (*Artificial Neural Network*, *Support Vector Machine*, dan *Random Forest*) untuk wilayah Kota Surabaya dengan menggunakan data citra satelit Landsat 8. Metode *Artificial Neural Network*, *Support Vector Machine*, dan *Random Forest* dipilih karena sebagaimana yang telah disebutkan, bahwa metode klasifikasi berbasis non parametrik yang sering digunakan adalah metode *Machine Learning Decision Trees*, *Random Forest*, *Artificial Neural Networks* dan *Support Vector Machines*, serta pada penelitian sebelumnya yang telah dilakukan oleh Harissalam (2018) telah dilakukan perbandingan klasifikasi tutupan lahan dengan metode *Random Forest* dan *Decision Trees*. Pada penelitian ini memiliki tujuan untuk mengevaluasi performa hasil klasifikasi yang dihasilkan dari ketiga metode tersebut (ANN, SVM, dan RF), sehingga perlu dilakukan analisis mengenai perbedaan hasil klasifikasi ketiga metode. Dimana, masing-masing metode tersebut

memiliki kelebihan masing-masing. Seperti untuk metode *Artificial Neural Network* memiliki keunggulan mampu menyelesaikan perhitungan paralel untuk tugas-tugas yang rumit, seperti prediksi dan pemodelan; klasifikasi dan pola pengenalan; pengklasteran; dan optimisasi (Pham, 1995). Untuk metode *Support Vector Machine* adapun keunggulan yang dimiliki diantaranya memiliki kemampuan yang lebih baik dalam menggeneralisasi data bila dibandingkan dengan teknik yang sudah ada sebelumnya (Vapnik dkk., 1997). Sedangkan, untuk metode *Random Forest*, yang merupakan metode yang terdiri dari gabungan (atau *forest*) dari beberapa *Trees*, memiliki keunggulan yakni mampu berjalan secara efisien untuk database yang besar dan mampu menghasilkan hasil perkiraan yang tidak bias dalam hasil *forestry* (pengelompokan) (Breiman, 2001). Bahasa Pemrograman R yang digunakan dalam penelitian ini memiliki beberapa kelebihan diantaranya bersifat *open source*, sehingga bisa digunakan oleh siapa saja secara gratis tanpa harus membayar lisensi. Selain itu, terdapat banyak *package* dan *function* di R yang memenuhi kebutuhan pengolahan data pada penelitian ini. Data primer yang digunakan berupa Citra Satelit Landsat 8 kanal OLI TIRS dengan resolusi spasial 30 meter. Dasar pemilihan data primer berupa citra satelit Landsat 8 adalah dikarenakan untuk melakukan analisis klasifikasi tutupan lahan, diperlukan data terkini (*up-to-date*) sehingga digunakan data primer berupa data citra satelit Landsat 8. Selain itu, penggunaan citra satelit Landsat 8 yang memiliki resolusi spasial menengah (30 meter) lebih efektif dan efisien dalam proses pengolahan dikarenakan tidak memerlukan perangkat keras dengan spesifikasi tinggi dan pengolahan membutuhkan waktu yang lebih cepat, mengingat Kota Surabaya sebagai area penelitian yang sangat besar dengan luas 32.680 ha. Kota Surabaya dipilih sebagai daerah penelitian karena daerah tersebut memenuhi semua jenis tutupan lahan yang ada, serta memiliki karakteristik fitur tutupan lahan yang kompleks dan cukup padat sehingga sesuai dengan tujuan tugas akhir dalam mengevaluasi akurasi dari metode klasifikasi tutupan lahan menggunakan pendekatan *Machine Learning Artificial Neural Network*, *Support Vector Machine*, dan *Random Forest*.

Data dan Metode

Surabaya secara geografis terletak pada $7^{\circ} 9'$ sampai dengan $7^{\circ} 21'$ Lintang Selatan dan $112^{\circ} 36'$ sampai dengan $112^{\circ} 54'$ Bujur Timur, secara umum kondisi topografi Kota Surabaya memiliki ketinggian tanah antara 0–20 meter di atas permukaan laut, sedangkan pada daerah pantai ketinggiannya berkisar antara 1–3 meter di atas permukaan laut. Luas Kota Surabaya adalah 274,06 km². Adapun batas Kota Surabaya adalah sebagai berikut :

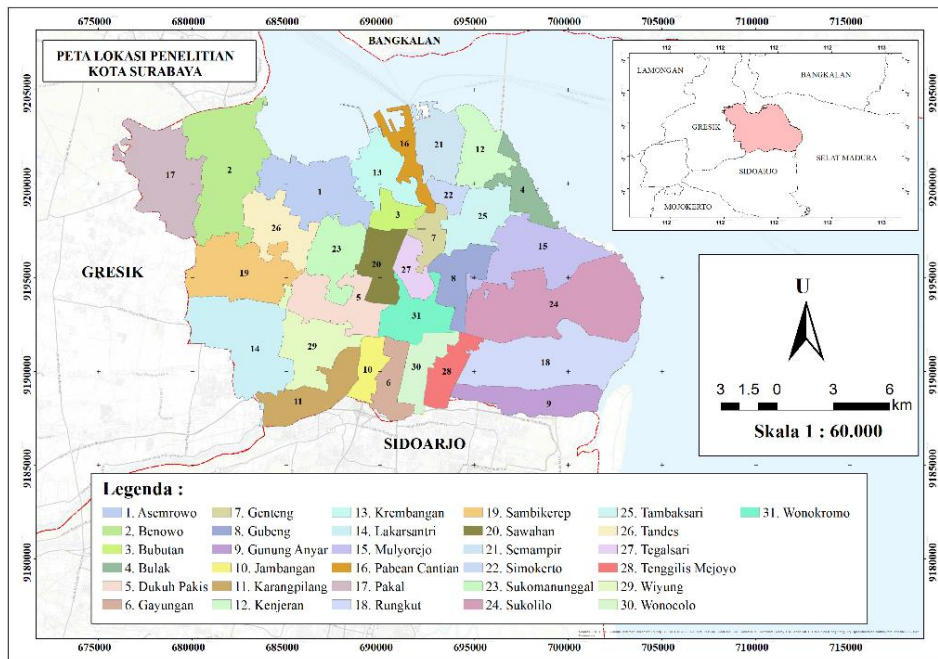
- Utara : Selat Madura
- Timur : Selat Madura
- Selatan : Kabupaten Sidoarjo
- Barat : Kabupaten Gresik

Data yang digunakan dalam penelitian ini adalah data Citra Satelit Landsat 8 OLI/TIRS C2 Level-1 wilayah Kota Surabaya tahun 2022 yang diakuisisi pada tanggal 2 Mei 2022 dengan resolusi spasial 30 meter yang didapatkan dari USGS. Satelit Landsat-8 memiliki sensor *Onboard Operational Land Imager* (OLI) dan *Thermal Infrared Sensor* (TIRS) dengan jumlah kanal sebanyak 11 buah. Diantara kanal-kanal tersebut, 9 kanal (band 1-9) berada pada OLI dan 2 lainnya (band 10 dan 11) pada TIRS. Sebagian besar kanal memiliki spesifikasi mirip dengan landsat 7. Tabel 1 menjelaskan karakteristik band-band yang terdapat pada citra Landsat 8.

Band yang digunakan pada penelitian ini hanya band 2, 3, 4, 5, 6, dan 7. Hal tersebut dikarenakan pada band 1 (panjang gelombang 0,43 - 0,45 μm) yaitu coastal aerosol dirancang untuk melakukan monitoring daerah perairan pesisir dan partikel halus, serta aerosol berkaitan erat dengan band 2 yaitu visible blue. Sedangkan untuk band Cirrus (panjang gelombang 1,36 – 1,38 μm) dirancang untuk identifikasi awan dan mengandung informasi permukaan daratan yang terbatas, sehingga tidak digunakan pada penelitian ini.

Proses pengolahan data pada penelitian ini terdiri atas beberapa proses. Pertama dilakukan pemilihan citra. Pada penelitian ini, digunakan data utama berupa citra satelit Landsat 8 C2/L1 OLI/TIRS yang diakuisisi pada tanggal 2 Mei 2022. Data citra satelit Landsat digunakan karena sangat sesuai dengan tujuan penelitian ini

yakni untuk melakukan klasifikasi tutupan lahan, yang membutuhkan data terbaru. Kemudian dilakukan *Pre-Processing* data citra. Citra satelit Landsat 8 yang digunakan memiliki resolusi spasial 30 meter dan harus dilakukan koreksi radiometrik dan atmosferik sehingga dapat digunakan dalam pengolahan data. Selanjutnya dilakukan pembuatan *Training Point*. Pada penelitian ini, pengambilan sampel dari data tutupan lahan yaitu 60 sampel untuk masing- masing kelas secara acak, hal tersebut dikarenakan, luas lokasi penelitian adalah 32.680 ha atau kurang dari 400.000 ha, dan jumlah kelas tutupan lahan yang digunakan untuk klasifikasi berjumlah 9 kelas atau kurang dari 12 kelas (Lillesand dan Kiefer, 1990). Adapun komposisi pembagian *training point* adalah 80% sebagai titik sample (*training data*) dan 20% komposisi *training point* sebagai *testing data*. *Training point* dibuat dengan format vektor (.shp), dan penempatan *training point* dilakukan sesuai dengan kelas klasifikasi yang telah ditentukan, dalam penelitian ini ada 4 kelas yaitu badan air, lahan terbangun, lahan terbuka, dan vegetasi.



Gambar 1. Lokasi Penelitian

Tabel 1. Band Citra Satelit Landsat 8 (USGS, 2013)

Band	Panjang gelombang (µm)	Sensor	Resolusi	Fungsi
1	0,43 - 0,45	Visible	30 m	Pesisir, partikel halus
2	0,45 - 0,51	Visible	30 m	Tutupan vegetasi
3	0,53 - 0,59	Visible	30 m	Vegetasi perbukitan
4	0,64 - 0,67	Near-infrared	30 m	Tutupan vegetasi
5	0,85 - 0,88	Near-infrared	30 m	Biomassa, garis pantai
6	1,57 - 1,65	SWIR 1	30 m	Kandungan air
7	2,11 - 2,29	SWIR 2	30 m	Kandungan air
8	0,50 - 0,68	Pankromatik	15 m	Mempertajam citra
9	1,36 - 1,38	Cirrus	30 m	Awan <i>cirrus</i>
10	10,6 11,19	TIRS 1	100 m	Pemetaan panas
11	11,5 - 12,51	TIRS 2	100 m	Pemetaan panas

Tahap selanjutnya adalah pembuatan *Training Machine Learning*. Data Training Sample Point merupakan sebaran titik sampel yang akan digunakan sebagai input data Machine Learning, untuk memasukkan nilai reflektan pada tiap titik sampel dilakukan melalui perintah R- Package (*raster::extract*). Proses klasifikasi dilakukan dengan menggunakan algoritma R, digunakan R-Package untuk memproses data *training sample point* yang telah didesain dan telah dilakukan uji akurasi model sebelumnya.

Pada hasil yang telah didapatkan, dilakukan uji akurasi. Uji akurasi dilakukan secara statistik melalui perangkat lunak RStudio dengan metode statistika *Confusion Matrix* (Koefisien Kappa, UA, dan PA) untuk melihat tingkat kesalahan yang terjadi pada klasifikasi yang dilakukan. Matriks konfusi dihitung dengan membandingkan hasil klasifikasi yang diturunkan dari citra dengan kondisi lapangan (*ground truth*) tutupan lahan. Setiap kolom dalam matriks konfusi mewakili kelas dari kondisi asli di lapangan (*ground truth*). Dan nilai dari kolom memiliki hubungan satu sama lain dengan label citra yang telah dipasang ke dalam *ground truth* pixel (Shao, dkk., 2016). Koefisien Kappa didapatkan melalui penghitungan statistik reliabilitas antar penilai, dan dihitung menggunakan algoritma sebagai berikut: (Zhenfeng S., dkk, 2016)

$$Kappa\ accuracy = \frac{N \sum_i^r x_{ii} - \sum_k^r x_{i+} x_{+i}}{N^2 - \sum_i^r x_{i+} x_{+i}} \quad (1)$$

Dimana:

- r = Jumlah baris dalam matriks konfusi
- x_{ii} = Jumlah pengamatan dalam baris i dan kolom i
- x_{i+} dan x_{+i} = Jumlah margin dari baris i dan kolom i
- N = Jumlah seluruh pengamatan

Producer's Accuracy (PA) merupakan probabilitas dari piksel yang terklasifikasi secara benar kedalam suatu tipe tutupan lahan, nilai *Producer's Accuracy* mewakili kesalahan akibat kelalaian. *User's Accuracy* (UA) merupakan proporsi dari piksel yang terklasifikasi secara benar, nilai UA mewakili dari kesalahan dalam pelaksanaan. Akurasi keseluruhan (*Overall Accuracy*) dihitung melalui perbandingan diantara jumlah piksel yang terklasifikasi secara benar dengan jumlah piksel yang digunakan dalam uji akurasi (Shao, dkk., 2016). Adapun persamaannya adalah sebagai berikut:

$$Producer's\ accuracy = \frac{x_{ii}}{x_{i+}} \times 100\% \quad (2)$$

$$User's\ accuracy = \frac{x_{ii}}{x_{+i}} \times 100\% \quad (3)$$

$$Overall\ accuracy = \frac{\sum x_{ii}}{N} \times 100\% \quad (4)$$

Berdasarkan nilai koefisien Kappa dilakukan klasifikasi nilai akurasi menjadi 6 kelas sebagai berikut:

Nilai Koefisien Kappa	Interpretasi Nilai Kappa
$0 <$	Peluang akurasi buruk
0,01 – 0,20	Peluang akurasi kecil
0,21 – 0,40	Peluang akurasi cukup
0,41 – 0,60	Peluang akurasi sedang
0,61 – 0,80	Peluang akurasi baik
0,81 – 0,99	Peluang akurasi sangat baik (<i>almost perfect</i>)

Analisis dilakukan dengan membandingkan hasil klasifikasi tutupan lahan dengan 3 metode *Machine Learning* yang berbeda yaitu *Artificial Neural Networks*, *Support Vector Machines*, dan *Random Forest*. Perbandingan dilakukan berdasarkan uji akurasi dengan menggunakan *confusion matrix*.

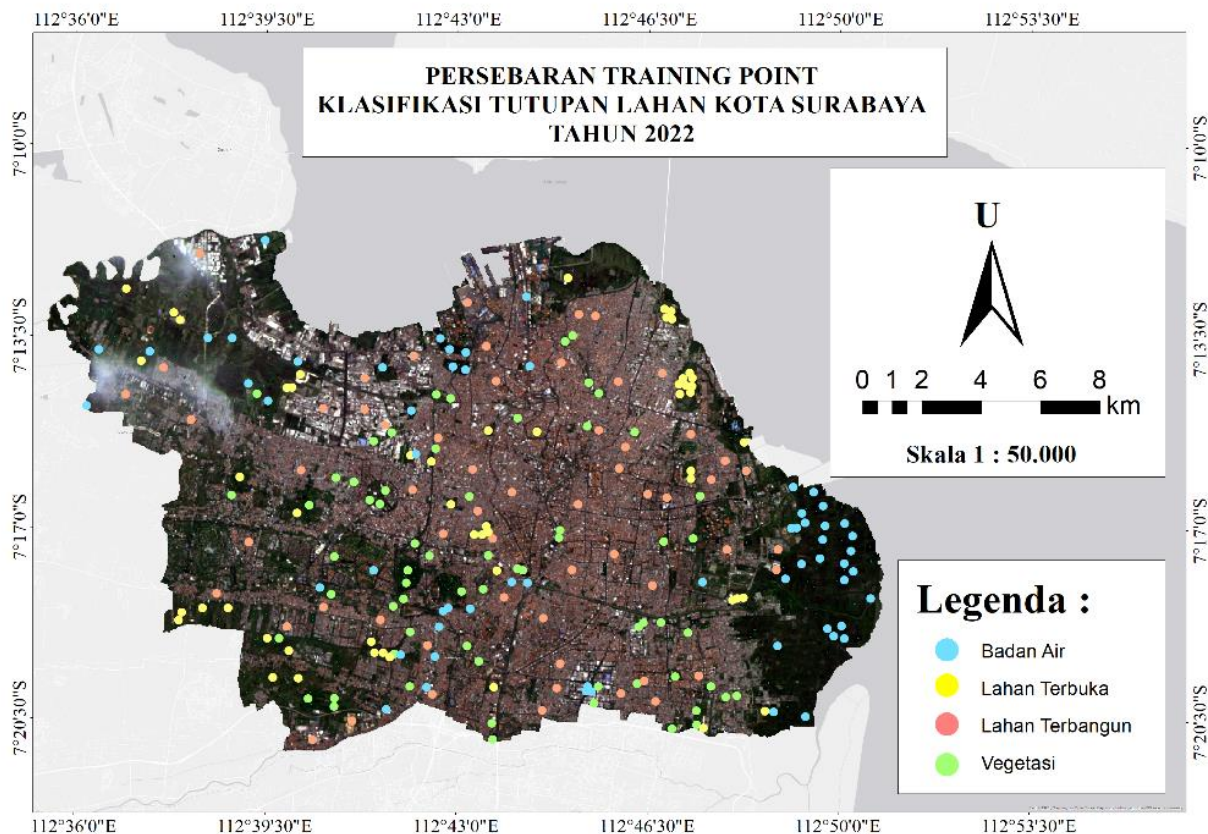
Hasil dan Pembahasan

Dalam penelitian ini, klasifikasi tutupan lahan dibagi atas 4 kelas, yaitu badan air, lahan terbuka, lahan terbangun, dan vegetasi. Yang termasuk ke dalam kelas badan air merupakan kumpulan air yang besar dan bentuknya antara lain bergantung pada relief permukaan bumi, seperti sungai, waduk, dan saluran. Untuk kelas

tutupan lahan terbuka merupakan area yang sengaja dibuka untuk masyarakat seperti area lapangan, makam, dan lahan terbuka non-hijau. Sedangkan lahan terbangun merupakan lahan yang sudah mengalami proses pembangunan atau pengerasan yang terjadi di atas lahan tersebut, yang meliputi area perumahan, perdagangan dan jasa, perkantoran, jalan, industri, sarana pelayanan umum. Tutupan lahan vegetasi merupakan kumpulan dari beberapa jenis tumbuhan yang tumbuh bersama-sama pada suatu tempat membentuk suatu kesatuan, seperti ruang terbuka hijau, pertanian (padi), dan taman. Sebelumnya telah dilakukan percobaan untuk menghasilkan klasifikasi tutupan dengan lebih banyak jumlah kelas tutupan lahan. Hasil klasifikasi tutupan lahan metode *machine learning* dengan menggunakan algoritma *Artificial Neural Network*, *Support Vector Machine*, dan *Random Forest* dilakukan uji akurasi yang menghasilkan nilai overall accuracy dan kappa accuracy. Hasil klasifikasi digunakan untuk menentukan metode dan algoritma *machine learning* yang paling baik diantara ketiga metode tersebut.

1. Training Point

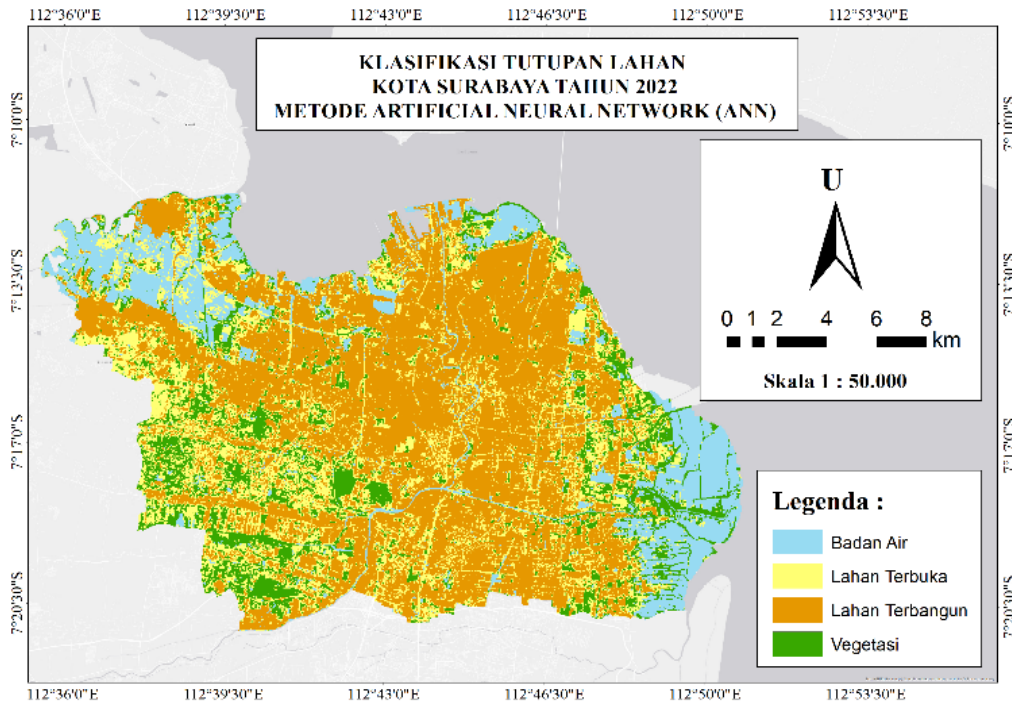
Training point dibuat dengan format vektor (.shp). Penempatan training point dilakukan dengan digitasi manual menggunakan prinsip interpretasi citra sesuai dengan kelas klasifikasi yang telah ditentukan. Dalam penelitian ini terdapat 4 kelas tutupan lahan (Badan Air, Lahan Terbuka, Lahan Terbangun, Vegetasi), dengan jumlah total *training point* sejumlah 237 titik dan tersebar sebagaimana tergambar pada Gambar 2. Penentuan komposisi training dan testing dilakukan secara acak oleh software R menggunakan algoritma **createDataPartition()**. Algoritma ini secara khusus akan digunakan untuk menemukan komposisi parameter model yang optimal dan untuk memeriksa kinerja model awal berdasarkan validasi silang berulang.



Gambar 2. *Training Point*

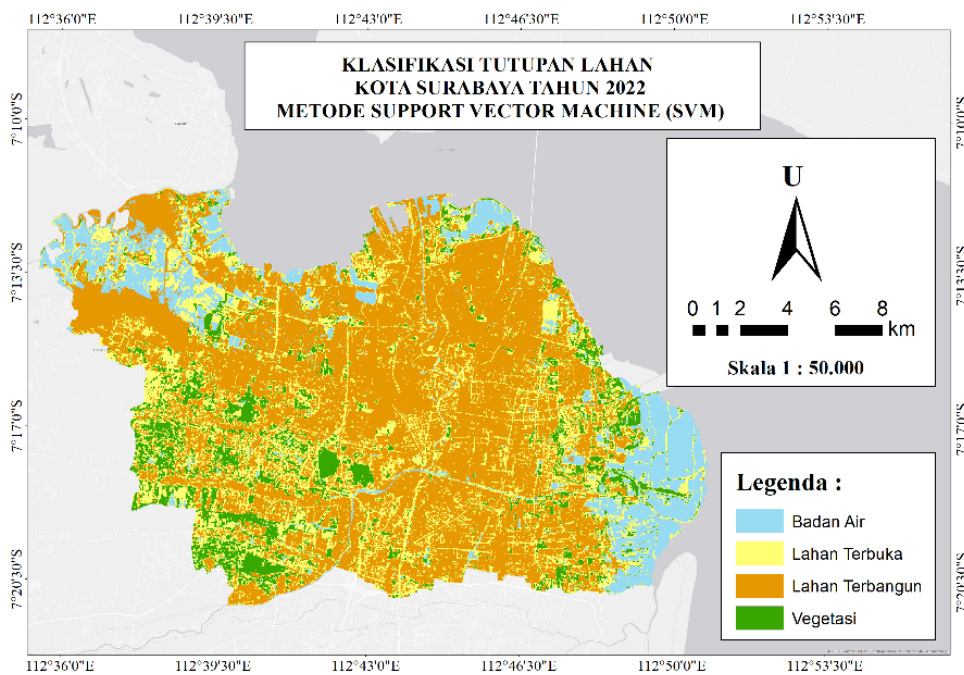
2. Hasil Klasifikasi Metode ANN

Klasifikasi tutupan lahan dari metode *Artificial Neural Network* menghasilkan peta tutupan lahan dengan nilai Overall Accuracy 91,11% dan nilai koefisien kappa 88,08%, dimana hasil ini didapatkan dari komposisi training point (80% titik sampel dan 20% titik validasi). Adapun hasil klasifikasi tutupan lahan, disajikan pada Gambar 3.



Gambar 3. Hasil Klasifikasi Tutupan Lahan dengan Metode ANN

3. Hasil Klasifikasi Metode SVM

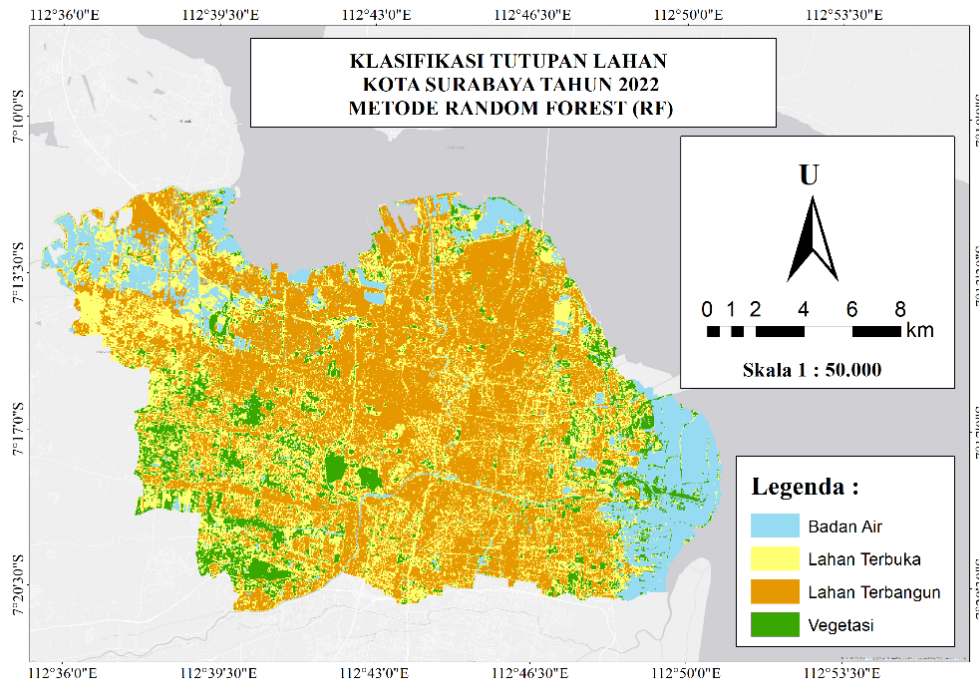


Gambar 4. Hasil Klasifikasi Tutupan Lahan dengan Metode SVM

Gambar 4 merupakan hasil klasifikasi tutupan lahan dengan metode SVM. Klasifikasi tutupan lahan dari metode *Support Vector Machine* menghasilkan peta tutupan lahan dengan nilai Overall Accuracy 91,11% dan nilai koefisien kappa 88,11%, dimana hasil ini dihasilkan dari komposisi training point (80% titik sampel dan 20% titik validasi).

4. Hasil Klasifikasi Metode RF

Klasifikasi tutupan lahan dari metode *Random Forest* menghasilkan peta tutupan lahan dengan nilai Overall Accuracy 93,33 % dan nilai koefisien kappa 91,07%, dimana hasil ini didapatkan dari komposisi training point (80% titik sampel dan 20% titik validasi). Adapun hasil klasifikasi tutupan lahan, disajikan dalam Gambar 5.



Gambar 5. Hasil Klasifikasi Tutupan Lahan dengan Metode RF

5. Uji Akurasi

Uji akurasi dilakukan secara statistik melalui perangkat lunak RStudio dengan metode statistika *Confusion Matrix* untuk melihat tingkat kesalahan yang terjadi pada klasifikasi yang dilakukan. Adapun hasil uji akurasi dengan menggunakan *confusion matrix* untuk masing-masing metode ditunjukkan pada Tabel 3, 4, dan 5.

Tabel 3. *Confusion matrix* metode ANN

Prediction	Reference				Total	Total Pixel Benar
	A	B	C	D		
A	12	1	0	0	13	41
B	0	6	0	0	6	
C	0	2	11	0	13	
D	0	1	0	12	13	
Total	12	10	11	12	45	
<i>Overall Accuracy</i>					0,9111	
<i>Kappa Accuracy</i>					0,8808	

Tabel 4. *Confusion matrix* metode SVM

Prediction	Reference				Total	Total Pixel Benar
	A	B	C	D		
A	12	1	0	0	13	42
B	0	6	0	0	6	
C	0	2	11	0	13	
D	0	1	0	12	13	
Total	12	10	11	12	45	
Overall Accuracy					0,9111	
Kappa Accuracy					0,8808	

Tabel 5. *Confusion matrix* metode RF

Prediction	Reference				Total	Total Pixel Benar
	A	B	C	D		
A	17	0	2	0	19	42
B	0	18	1	0	19	
C	0	0	10	0	10	
D	1	0	2	13	16	
Total	18	18	13	13	64	
Overall Accuracy					0,9333	
Kappa Accuracy					0,9107	

Keterangan :

- (A) Badan Air
- (B) Lahan Terbuka
- (C) Lahan Terbangun
- (D) Vegetasi

Berdasarkan Tabel 4, pada metode ANN dari 45 titik testing yang digunakan untuk validasi, 41 titik dianggap benar. Sedangkan untuk Tabel 5, pada metode SVM dari 45 titik testing yang digunakan untuk validasi, 41 titik dianggap benar. Dan pada Tabel 6, dari 45 titik testing yang digunakan untuk validasi, 42 titik dianggap benar. Nilai *overall accuracy* yang didapatkan pada ketiga metode sudah memenuhi standar yang ditetapkan oleh *United States Geological Survey* (USGS) yakni berada pada nilai di atas 85%. Sedangkan untuk *kappa accuracy* berdasarkan Koefisien Kappa Cohen seluruh metode masuk dalam kategori peluang akurasi sangat baik (*almost perfect*).

Tabel 6. Nilai *User's Accuracy* dan *Producer's Accuracy* Metode ANN

Kelas Tutupan Lahan	UA	PA
Badan Air	94,11	88,89
Lahan Terbuka	100,00	100,00
Lahan Terbangun	80,59	80,00
Vegetasi	92,11	89,74

Tabel 7. Nilai *User's Accuracy* dan *Producer's Accuracy* Metode SVM

Kelas Tutupan Lahan	UA	PA
Badan Air	94,44	94,44
Lahan Terbuka	90,00	100,00
Lahan Terbangun	77,78	76,66
Vegetasi	81,39	89,74

Tabel 8. Nilai *User's Accuracy* dan *Producer's Accuracy* Metode RF

Kelas Tutupan Lahan	UA	PA
Badan Air	89,47	94,44
Lahan Terbuka	94,74	100,00
Lahan Terbangun	100,00	86,66
Vegetasi	92,85	100,00

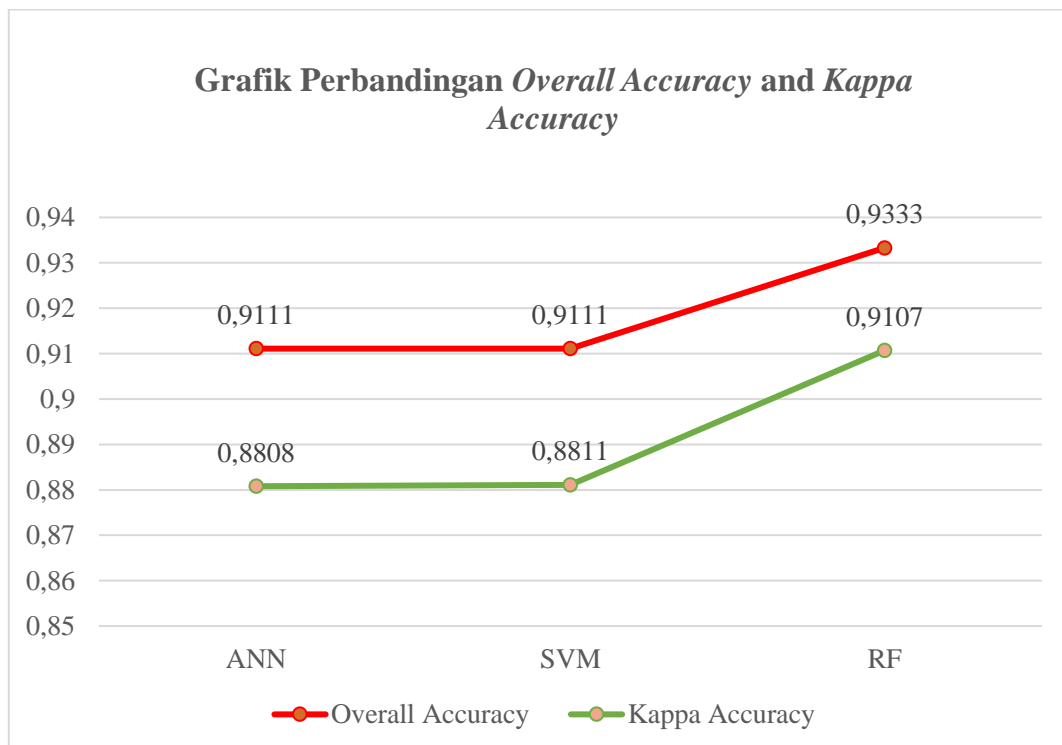
Keterangan:

UA: *User's Accuracy*

PA: *Producer's Accuracy*

Producer's Accuracy merupakan peluang rata-rata (%) suatu piksel akan diklasifikasikan dengan benar dan secara rata-rata menunjukkan seberapa baik setiap kelas di lapangan telah diklasifikasi. Sedangkan *User's Accuracy* merupakan peluang rata-rata (%) suatu piksel dari citra yang telah terklasifikasi, secara aktual mewakili kelas-kelas tersebut di lapangan. Berdasarkan Tabel 7, 8, dan 9 dihasilkan *user's accuracy* untuk tiap tutupan lahan bernilai diatas 85%, kecuali pada kelas lahan terbangun (metode SVM dan RF). Hal tersebut, terjadi dikarenakan pada metode ini terjadi kesalahan klasifikasi lahan terbangun ke dalam kelas lahan terbuka.

6. Perbandingan Hasil Klasifikasi



Gambar 6. Perbandingan Hasil Overall Accuracy dan Kappa Accuracy Klasifikasi Tutupan Lahan dengan Metode ANN, SVM, dan RF

Berdasarkan hasil yang telah didapatkan dan uji akurasi yang telah dilakukan, nilai akurasi yang didapatkan dari ketiga metode menunjukkan hasil yang baik menurut ketentuan yang ditetapkan oleh *United States Geological Survey* (USGS) dengan tingkat ketelitian klasifikasi atau *overall accuracy* sebesar 85%. Sedangkan untuk nilai kappa accuracy berdasarkan Koefisien Kappa Cohen masuk dalam kategori Peluang akurasi sangat baik (*almost perfect*). Dari ketiga metode *machine learning*, algoritma *Random Forest* memberikan hasil yang paling baik, dengan nilai *Overall Accuracy* 93,33 % dan nilai koefisien kappa 90,41%. Sedangkan metode *Artificial Neural Network* merupakan algoritma terbaik kedua dengan *Overall Accuracy* 90,00 % dan nilai koefisien kappa 87,87 %. Dan metode, memiliki nilai akurasi yang paling rendah bila dibandingkan dengan ketiga metode yang lain adalah metode *Support Vector Machine*, dengan nilai *Overall accuracy* 87,56 % dan nilai koefisien kappa 85,12%.

Kesimpulan

Pada penelitian ini, dilakukan klasifikasi tutupan lahan dengan menggunakan metode *machine learning* dengan menggunakan bahasa pemrograman R dan data citra Satelit Landsat 8 C2/L1 OLI/TIRS untuk Kota Surabaya. Adapun perbandingan komposisi training point (80% titik sampel dan 20% titik validasi). Adapun metode yang digunakan dalam penelitian ini adalah metode *Artificial Neural Network*, *Support Vector Machine*, dan *Random Forest*, dengan menggunakan bahasa pemrograman R. Pada penelitian ini, dihasilkan klasifikasi tutupan lahan dengan klasifikasi tutupan lahan dibagi atas 4 kelas, yaitu badan air, lahan terbuka, lahan terbangun, dan vegetasi.

Dari ketiga metode *machine learning*, *Random Forest* merupakan metode yang paling baik dengan nilai *overall accuracy* sebesar 93,33% dan *kappa accuracy* sebesar 91,07%. Sedangkan metode *Support Vector Machine* merupakan metode terbaik kedua dengan nilai *overall accuracy* sebesar 91,11% dan *kappa accuracy* sebesar 88,11%. Dan metode *Artificial Neural Network* merupakan metode yang memiliki nilai paling rendah dibandingkan dengan metode yang lain dengan nilai *overall accuracy* sebesar 91,11% dan *kappa accuracy* sebesar 88,08%.

Apabila dilakukan penelitian serupa dengan menggunakan metode *machine learning* yang dimanfaatkan untuk klasifikasi tutupan lahan, penulis menyarankan untuk menggunakan data citra satelit resolusi tinggi agar lebih mudah dalam melakukan training point dan hasil perbandingan klasifikasi tutupan lahan yang didapatkan lebih akurat. Selain itu, perlu diperhatikan pula dalam menentukan titik training.

Ucapan Terimakasih

Penulis mengucapkan terima kasih kepada Dinas Cipta Karya dan Tata Ruang Kota Surabaya yang telah menyediakan data guna menunjang penelitian ini dan juga kepada kerabat yang sudah membantu dalam penelitian ini. terima kasih kepada sponsor penyedia dana penelitian.

Daftar Pustaka

- Breiman, L. (2001). *Random Forest*. California : Handbook of Statistics Department University of California.
- Harissalam, M., (2020). Perbandingan Hasil Klasifikasi Tutupan Lahan Pada Citra Satelit Resolusi Tinggi Pleiades Menggunakan Pendekatan Machine Learning Metode Decision Tree Dan Random Forest (Studi Kasus : Kelurahan Kejawan Putih Tambak, Kecamatan Mulyorejo, Surabaya). Undergraduate Thesis, Dept. Geomatics Eng, Institut Teknologi Sepuluh Nopember (2020).
- Havyas, V. B., dkk. (2015). Decision Tree Approach for Classification of Satellite Imagery. *Int. J. Soft Computing Eng.(IJSCE)*, 5(2), 2231-2307.
- Jensen, J. R., & Cowen, D. C. (1999). Remote sensing of urban/suburban infrastructure and socio-economic attributes. *Photogrammetric engineering and remote sensing*, 65, 611-622.
- Lillesand, T. M. & Kiefer, R. W. (1990). *Penginderaan Jauh dan Interpretasi Citra Alih Bahasa R*. Yogyakarta : Universitas Gadjah Mada.

- Nugroho, R. A., & Handayani, H. H. (2021). Prediksi Perubahan Tutupan Lahan Menggunakan Metode Markov Chain dan Citra Satelit Penginderaan Jauh (Studi Kasus: Kota Surabaya). *Jurnal Teknik ITS*, 9(2), C71-C77.
- Pham, D.T. (1994). *Neural Network for Chemical Engineers*. Amsterdam : Elsevier Press.
- RStudio. (2021). *About RStudio*. [Online]. Available : <https://www.rstudio.com/about/>.
- Shao, Z., Fu, H., Fu, P., & Yin, L. (2016). Mapping Urban Impervious Surface by Fusing Optical and SAR Sata at the Decision Level. *Remote Sensing*, 8(11), 94.
- Suryo, A. B., & Hariyanto, T. (2013). Studi Perubahan Tutupan Lahan DAS Ciliwung Dengan Metode Klasifikasi Terbimbing Citra Landsat 7 ETM+ Multitemporal Tahun 2001 & 2008 (Studi Kasus: Bogor). *Geoid*, 9(1), 1-6.
- Natural Resouces Canada. (2015). [Online]. *Land Use Mapping*. Available : <https://www.nrcan.gc.ca/>
- Vapnik, V., Golowich, S. E., & Smola, A. (1997). Support Vector Method for Function Approximation, Regression Estimation and Signal Processing. *Advances in Neural Information Processing Systems 9*.



This article is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/).