

Deteksi Tingkat Kepadatan Rumah Penduduk Menggunakan Klasifikasi K-Nearest Neighbor Pada Citra Satelit Google Earth

Abd. Ghofur

Universitas Ibrahimy
Email: apunkbwi@gmail.com

(Naskah masuk: 25 Juni 2020, diterima untuk diterbitkan: 8 Juli 2020)

ABSTRAK

Penelitian ini berupa perancangan dan Implementasi computer vision berbasis digital image processing pada citra yang diambil dari satelit google earth dan bertujuan mengetahui kepadatan rumah penduduk dan lahan terbuka suatu wilayah sebagai deteksi awal. Penelitian ini juga bertujuan untuk memberikan kontribusi keilmuan di bidang digital image processing sebagai alternatif usulan model deteksi kepadatan rumah penduduk.

Metodology yang digunakan dalam penelitian ini adalah studi kepustakaan yang diperoleh dari jurnal yang diidentik. Proses yang dilakukan adalah mengambil citra dari suatu desa atau daerah pada google earth yang berbentuk RGB, mengekstraksi nilai pixel RGB, mendapatkan nilai mean RGB, mentransformasikan RGB ke HSI dan mencari nilai mean HSI, mencari variance dan masing-masing RGB dan HSI kemudian diklasifikasi menggunakan K-Nearest Neighbor.

Dataset diklasifikasi menggunakan K-Nearest Neighbor menjadi tiga kategori yaitu rumah penduduk yang padat, sedang dan longgar. Selanjutnya dataset yang telah diklasifikasi dijadikan data training sebagai perbandingan data untuk testing. Hasil testing menunjukkan bahwa akurasi deteksi kepadatan rumah penduduk dari citra testing dengan membanding hasil pengamatan ahli mencapai 95%.

Kata kunci: *Deteksi, Kepadatan Rumah, Penduduk, K-NN, RGB, HSI*

ABSTRACT

This research is in the form of design and implementation of computer vision based on digital image processing on images taken from Google Earth satellite and aims to determine the density of people's homes and open land of an area as an early detection. This study also aims to make scientific contributions in the field of digital image processing as an alternative to the proposed model for the detection of population densities.

The methodology used in this research is literature study obtained from identified journals. The process is taking an image from a village or region on Google earth in the form of RGB, extracting the RGB pixel value, getting the RGB mean value, transforming RGB to HSI and looking for the mean value of HSI, looking for variance and respectively RGB and HSI then classified using K-Nearest Neighbor.

The dataset is classified using K-Nearest Neighbor into three categories, namely dense, medium and loose residents' houses. Furthermore, the dataset that has been classified into training data is used as a comparison of data for testing. The results of testing showed that the accuracy of detection of population densities of houses from testing images by comparing the results of expert observations reached 95%.

Keywords: *Detection, House Density, Population, K-NN, RGB, HSI*

1. PENDAHULUAN

Kepadatan rumah penduduk setiap tahun semakin bertambah seiring dengan meningkatnya pertumbuhan penduduk dan

ekonomi, sementara data kepadatan penduduk membutuhkan waktu lama yang diambil dari sensus penduduk. Sementara penelitian lain menyebutkan bahwa jumlah penduduk, pendidikan dan pengangguran

secara bersama-sama berpengaruh terhadap kemiskinan. Jumlah penduduk yang besar disertai sumberdaya manusia yang bermutu, maka semakin tinggi produktivitas ekonomi sehingga setiap keluarga baru membutuhkan rumah baru. (Mahsunah, 2013)

Informasi jumlah rumah penduduk didapat dari sensus pendudukan yang dilakukan oleh bapenas bahwa proyeksi penduduk indonesia tahun 2010-2035 diasumsikan pada level fertilitas 1,2 anak perwanita. (Kementrian PPN/Bappenas, 2013)

Dengan asumsi tersebut maka sensus rumah penduduk dituntut untuk lebih cepat dan akurat sehingga pemerintah berinovasi dengan menggunakan sensus online.

Berdasarkan beberapa hal di atas, maka peneliti mencoba untuk melakukan penelitian sebagai usulan alternatif model bagi pemerintah atau pihak terkait sebagai deteksi awal kepadatan rumah di suatu daerah.

Batasan dalam penelitian ini adalah sebagai berikut :

- a. Penelitian menggunakan studi kepustakaan yang didapatkan dari buku, jurnal dan browsing.
- b. Citra yang diperoleh dari citra satelit google eart dengan jarak + 1000 m.
- c. Wilayah yang dijadikan citra sample adalah wilayah jawa timur (daerah beriklim tropis) .

Deteksi kepadatan rumah penduduk merupakan usaha menemukan dan menentukan keberadaan, anggapan, atau kenyataan yang berkaitan dengan rumah penduduk. (Kamus Besar Bahasa Indonesia, 2016) Sedangkan kepadatan rumah penduduk bedasarkan Badan Standar Nasional, tentang tata cara perencanaan lingkungan perumahan di perkotaan SNI 03-1733-2004 menjelaskn bahwa kepadatan lingkungan rata-rata 50 rumah per hektar dan rata-rata jumlah

keluarga 4-5 jiwa/keluarga. (Badan Standar Nasional, 2004)

Pengolahan citra merupakan proses pengolahan dan analisis pada sebuah citra yang banyak melibatkan persepsi visual. Pada proses ini mempunyai ciri data masukan informasi keluaran yang berbentuk citra. (Yousman and Yeyep, 2018) Istilah pengolahan citra digital secara umum didefinisikan sebagai pemrosesan citra dua dimensi dengan komputer.

Google Earth merupakan aplikasi pemetaan yang dikeluarkan oleh Google Corp. Google Earth menampilkan peta bola dunia, keadaan topografi, terrain yang dapat di-overlay dengan jalan, bangunan lokasi ataupun informasi geografis lainnya. (Sarie et al., 2016)

Citra RGB merupakan citra yang menggunakan basis tiga warna yaitu warna Red (merah), Green (hijau) dan Blue (biru). Dari kombinasi ketiga warna tersebut dapat mewakili warna objek seperti warna aslinya.

HSI (Hue, Saturasi dan Intensitas) yang dikenal dengan 3 dimensi ruang. Hue merupakan warna dasar yang meliputi merah, kuning, hijau dan biru atau campuran dari warna tersebut. Saturasi yaitu ketajaman warna dari hue. Sedangkan Saturasi merupakan pencahayaan dari hue dan saturasi. (Saputra and Arifin, 2017)

Ekstraksi ciri citra merupakan proses pencarian informasi dari citra digital dimana informasi tersebut digunakan sebagai langkah awal untuk proses klasifikasi. (Surya, 2016)

Algoritma K-Nearest Neighbor (K-NN) merupakan metode klasifikasi yang digunakan untuk menentukan label dari sebuah objek baru berdasarkan mayoritas label dari jarak K dari dari kelompok beberapa data latih. (Yustanti, 2012)

Dari beberapa kajian teoritis di atas maka penelitian ini memiliki tujuan yaitu
p-ISSN : 2502-5724; e-ISSN : 2541-5735

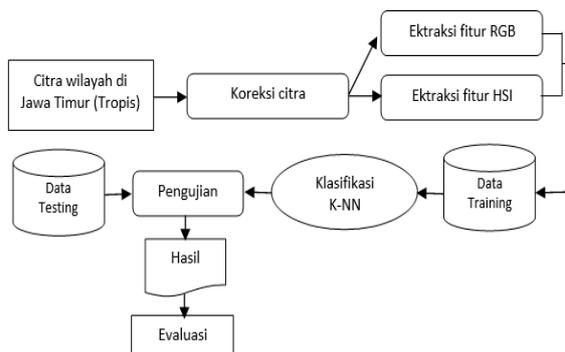
mengusulkan alternatif model sebagai deteksi awal kepadatan penduduk suatu wilayah di Jawa Timur.

2. METODOLOGI

Penelitian menggunakan pendekatan metode deskriptif analitik melalui perancangan piranti lunak (software).

Alat yang digunakan adalah komputer yang sudah dipasang software aplikasi google earth untuk mendapatkan citra wilayah wilayah tertentu di Jawa Timur yang memiliki fitur RGB. Citra dari google earth inilah yang menjadi bahan untuk diproses selanjutnya dengan beberapa metode pada image processing.

Peneitian ini terbagi dalam beberapa tahap. Tahapan tersebut dapat dilihat pada gambar 1 berikut :



Gambar 1. Tahapan Penelitian

Tahapan dalam metode pelaksanaan penelitian yaitu sebagai berikut :

- Persiapan dan studi kepustakaan
- Pada tahap ini melakukan persiapan dan perencanaan kemudian melakukan studi kepustakaan tentang literatur idektik dengan topik penelitian yang dibutuhkan.
- Pengambilan citra digital dari citra satelit google eart dan koreksi citra dengn menyeragamkan ukuran yaitu 1024 x 720.
- Mengekstraksi RGB dan tranformasi ke HSI

- Mengklasifikasi dataset dari citra.
- Mengambil data dari dataset menjadi data training sebanyak 80 data.
- Pengujian data testing sebanyak 20 data.
- Informasi hasil pengujian
- Mengevaluasi hasil pengujian berdasarkan pengmatan ahli.

3. HASIL DAN PEMBAHASAN

Pada tahap ini menjelaskan hasil penelitian dan pembahasan yang terdiri dari beberapa tahap, yaitu:

3.1. Penyiapan data dan koreksi citra

Tahap berikutnya adalah mempersiapkan data yang memanfaatkan aplikasi google earth, yaitu dengan memiliki wilayah desa atau kelurahan tertentu dengan jarak +1000 m dan menyimpan citra dengan format jpg. Data yang disiapkan sebanyak 100 citra yang akan dibagi dalam dua kelompok yaitu 80 citra untuk data training dan 20 citra untuk data testing.

Contoh citra dataset citra rumah penduduk longgar dapat dilihat pada Gambar 2.



Gambar 2. Dataset Citra Untuk Rumah Penduduk Longgar

Dataset citra untuk rumah penduduk sedang dapat dilihat pada gambar 3.



Gambar 3. Dataset Citra Untuk Rumah Penduduk Sedang

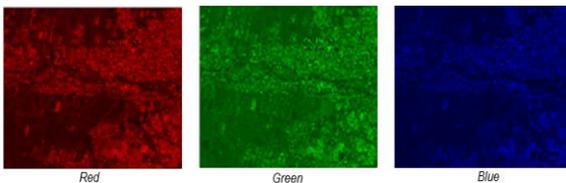
Sedangkan dataset citra untuk rumah penduduk yang pada dapat dilihat pada gambar 4.



Gambar 4. Dataset Citra Untuk Rumah Penduduk Sedang

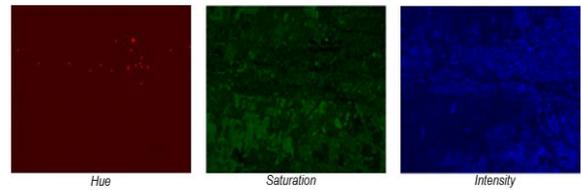
3.2. Ekstraksi fitur RGB dan HSI

Tahap ini mengekstraksi citra RGB dan mentransformasikan citra RGB ke HSI. Citra RGB dan HSI. Citra RGB dapat dilihat pada gambar 5.



Gambar 5. Citra RGB

Sedangkan citra HSI dapat dilihat pada gambar 6.



Gambar 6. Citra HSI

Ekstraksi RGB bertujuan untuk mendapat informasi berupa nilai mean, variance, range dan standar deviasi dari citra Red, Green dan blue. Selanjutnya mentransformasikan citra RGB ke HSI bertujuan untuk mendapatkan informasi berupa nilai mean, variance, range dan standar deviasi dari citra Hue, Saturation dan Intensity. Transformasi RGB ke HSI menggunakan rumus :

Menghitung θ

$$\theta = \cos^{-1} \left\{ \frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\} \quad (1)$$

Menghitung H (Hue)

$$H = \begin{cases} \theta & \text{jika } B \leq G \\ 360 - \theta & \text{jika } B > G \end{cases} \quad (2)$$

jika $B \leq G$
jika $B > G$

Menghitung S (Saturation)

$$S = 1 - 3 \frac{\min(R,G,B)}{(R+G+B)} \quad (3)$$

Menghitung I (Intensity)

$$I = \frac{1}{3} (R + G + B) \quad (4)$$

Selanjutnya menghitung range dari masing-masing sampel RGB dan HSI dengan rumus :

$$R = x_{max} - x_{min} \quad (5)$$

Dimana R adalah rentang atau jangkauan, x_{max} adalah data yang paling besar sedangkan x_{min} adalah data yang paling kecil.

Langkah selanjutnya adalah menghitung varians dari sampel RGB dan HSI dengan rumus :

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n-1} \tag{6}$$

dimana :

S^2 = sampel varian

x_i = nilai dari satu observasi

\bar{x} = nilai dari seluruh observasi

n = jumlah observasi

Kemudian menghitung standar deviasi dengan rumus:

$$S = \sqrt{\frac{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)}} \tag{7}$$

dimana :

s = standar deviasi

x_i = nilai x ke i

n = ukuran sampel

Selanjutnya menyimpan informasi data fitur citra yang telah diekstraksi. Informasi fitur yang tersimpan tersebut digunakan sebagai data training.

Adapun data training yang telah diekstraksi fitur dengan diperoleh informasi berupa mean red, mean green, Mean Blue, Mean Hue, mean Saturasi dan mean Intensitas dapat dilihat pada tabel 1.

Tabel 1. Ekstraksi Mean RGB dan HSI

NO	MeanR	MeanG	MeanB	MeanH	MeanS	MeanI
1	0,4021455	0,3684242	0,3495852	0,141134	0,1114068	0,3733849
2	0,3502765	0,3645226	0,3037029	0,2062785	0,1324915	0,3395006
3	0,360939	0,3693653	0,3559725	0,239718	0,0672896	0,3620923
4	0,3722622	0,3721109	0,3173627	0,1824331	0,1349913	0,3539119
5	0,4765189	0,4776161	0,4705316	0,2329546	0,0743338	0,4748888
...						
22	0,3289237	0,3484636	0,2810016	0,2094252	0,1339562	0,319463
23	0,4220191	0,4665835	0,3965398	0,260955	0,0907073	0,4283808
24	0,5393854	0,5443132	0,4790443	0,1862912	0,0844384	0,5209143
25	0,3746907	0,4030305	0,3330437	0,2227946	0,1195689	0,370255
...						
77	0,1200779	0,1464783	0,121988	0,3433015	0,1048084	0,1295147
78	0,370998	0,4181962	0,3448901	0,2679334	0,1322784	0,3780281
79	0,1469148	0,169466	0,1339115	0,2742523	0,1284075	0,1500974
80	0,2839379	0,3022597	0,2504559	0,2380773	0,1342146	0,2788845

Data training yang telah diekstraksi fitur dengan diperoleh informasi berupa variance red, variance green, variance

Blue, variance Hue, variance Saturasi dan variance Intensitas dapat dilihat pada tabel 2.

Tabel 2. Ekstraksi Variance RGB dan HSI

NO	VarRed	VarGreen	VarBlue	VarH	VarS	VarI
1	0,031219	0,0214191	0,0208426	0,0243863	0,0090165	0,0234283
2	0,0213476	0,012951	0,014294	0,0124126	0,0049232	0,0155057
3	0,0236558	0,0192172	0,018744	0,0274722	0,0069615	0,0201503
4	0,027031	0,0162356	0,017589	0,0151451	0,0054832	0,019321
5	0,0212235	0,0124337	0,0118351	0,0267176	0,0044516	0,0140496
...						
22	0,0168794	0,0104806	0,0111334	0,0092478	0,0023835	0,0122837
23	0,0114087	0,0060727	0,0066685	0,0075768	0,0012303	0,0076177
24	0,0153594	0,00838	0,0094206	0,0090612	0,0007998	0,0105094
25	0,023431	0,0127958	0,0160073	0,0113469	0,0024181	0,016651
...						
77	0,0016648	0,00185	0,0012603	0,0054622	0,0022848	0,0014975
78	0,014933	0,0071367	0,0084555	0,0149318	0,0028652	0,0090812
79	0,0033124	0,0033178	0,0025652	0,00744	0,0045853	0,0029165
80	0,0118594	0,0056399	0,0050462	0,0134234	0,0036758	0,0069329

Data training yang telah diekstraksi fitur dengan diperoleh informasi berupa range red, range green, range Blue, range Hue, range Saturasi dan range Intensitas dapat dilihat pada tabel 3.

Tabel 3. Ekstraksi Range RGB dan HSI

NO	RangeR	RangeG	RangeB	RangeH	RangeS	RangeI
1	1	1	1	0,5	1	0,9973856
2	1	1	1	0,5	1	0,9908497
3	1	1	1	0,5	1	1
4	1	1	1	0,5	1	1
5	1	1	1	0,5	1	0,9973856
...						
22	1	1	1	0,5	1	0,9921569
23	1	1	0,9921569	0,5	1	0,9895425
24	1	0,9921569	0,9960784	0,5	1	0,9882353
25	1	1	1	0,5	1	0,9908497
...						
77	0,9921569	1	1	0,5	1	0,9960784
78	1	1	1	0,5	1	0,9947712
79	1	1	1	0,5	1	1
80	1	1	0,972549	0,5	1	0,9673203

Data training yang telah diekstraksi fitur dengan diperoleh informasi berupa standar deviasi red, standar deviasi green, standar deviasi Blue, standar deviasi Hue, standar deviasi Saturasi dan standar deviasi Intensitas dapat dilihat pada tabel 4.

Tabel 4. Ekstraksi Range RGB dan HSI

NO	SdR	SdG	SdB	SdH	SdS	SdI
1	0,176689	0,1463527	0,1443698	0,1561611	0,0949552	0,1530631
2	0,1461081	0,1138025	0,1195574	0,1114119	0,0701658	0,1245219
3	0,1538046	0,1386262	0,1369086	0,1657475	0,0834356	0,1419518
4	0,1644111	0,1274189	0,1326236	0,1230654	0,0740486	0,1389999
5	0,1456827	0,1115067	0,1087892	0,1634552	0,0667207	0,1185311
...						
22	0,1299208	0,1023747	0,1055147	0,0961653	0,0488211	0,1108317
23	0,1068115	0,0779275	0,0816611	0,087045	0,035076	0,0872792
24	0,1239331	0,0915426	0,0970598	0,0951903	0,0282803	0,1025155
25	0,1530718	0,1131183	0,1265198	0,106522	0,0491738	0,1290387
...						
77	0,0408022	0,0430111	0,0355011	0,0739065	0,0477994	0,0386974
78	0,1222006	0,0844792	0,0919537	0,1221958	0,0535276	0,0952955
79	0,0575539	0,0576003	0,0506478	0,0862556	0,067715	0,0540049
80	0,108901	0,0750996	0,0710364	0,1158593	0,0606287	0,0832637

Berdasarkan tabel 1,2,3 dan 4 di atas, data dikategorikan menjadi tiga macam. Kategori 1 menunjukkan bahwa data yang berarti tingkat kepadatan rumah penduduk

“Longgar”, kategori 2 menunjukkan bahwa data yang tingkat kepadatan rumah penduduk “Sedang” dan kategori 3 menunjukkan bahwa data yang tingkat kepadatan rumah penduduk “Padat”.

3.3. Pengujian

Pada tahap ini dilakukan pengujian untuk citra baru. Tujuan pengujian adalah mengetahui tingkat akurasi atau ketepatan dalam mendeteksi citra dalam tiga kategori longgar, sedang dan padat.

Tahap ini mengklasifikasi data training yang telah di ekstraksi dari citra RGB dan HSI. Metode klasifikasi menggunakan metode K-Nearest Neighbor. Metode klasifikasi metode K-Nearest Neighbor ini dipilih pada penelitian ini karena memilih beberapa keunggulan diantaranya mampu mengklasifikasi terhadap data yang memiliki banyak noise dan efektif untuk data latih yang besar. (Wijaya and Ridwan, 2019)

Klasifikasi pada dataset dibagi dalam tiga kategori, yaitu yaitu kepadatan rumah penduduk kategori padat, sedangkan dan longgar.

Untuk mengklasifikasi dengan metode K-Nearest Neighbor menggunakan rumus (Yustanti, 2012) :

$$distance = \sqrt{\sum_{i=1}^n (X_{training}^i - X_{testing})^2} \quad (8)$$

dengan

- $X_{training}^i$: data training ke-i,
- $X_{testing}^i$: data testing
- i : record (baris) ke-i dari tabel,
- n : jumlah data training.

Citra yang digunakan untuk pengujian ini sebanyak 20 citra.

Pada citra uji juga diekstraksi RGB dan ditransformasi ke HSI untuk mendapatkan nilai mean, variance, range dan standar deviasi dari citra red, green dan blue begitu juga dengan ciutra hue, saturation dan intensity.

Setelah beberapa fitur didapatkan dari RGB dan HSI, maka selanjutnya data

testing diuji menggunakan metode K-Nearest Neighbor dan ditentukan hasilnya, apakah kepadatan rumah penduduk berdasarkan citra termasuk kategori longgar, sedang atau padat.

Data hasil pengujian dapat dilihat pada tabel 5.

Tabel 5. Hasil Pengujian

No	Nama Image Testing	Aplikasi	Pengamatan	Keterangan
1	Binakal	LONGGAR	LONGGAR	BENAR
2	Bungatan	LONGGAR	LONGGAR	BENAR
3	Gambiran	LONGGAR	LONGGAR	BENAR
4	Glenmore	LONGGAR	LONGGAR	BENAR
5	Licin	LONGGAR	LONGGAR	BENAR
...
13	Ngawi	LONGGAR	SEDANG	SALAH
14	Situbondo	SEDANG	SEDANG	BENAR
15	Srono	SEDANG	SEDANG	BENAR
16	Mojokerto	PADAT	PADAT	BENAR
17	Pamekasan	PADAT	PADAT	BENAR
18	Probolinggo	PADAT	PADAT	BENAR
19	Sampang	PADAT	PADAT	BENAR
20	Surabaya	PADAT	PADAT	BENAR

3.4. Evaluasi

Hasil evaluasi dari 20 citra uji bahwa terdapat satu citra yang tidak dapat dideteksi dengan oleh sesuai dengan pengamatan ahli sehingga hal ini dapat mengurangi akurasi ketepatan dektesi. Dari pengamatan ahli bahwa citra tersebut termasuk kategori sedang, sedangkan aplikasi mengkategorikan citra tersebut termasuk pada kategori longgar.

4. KESIMPULAN

Berdasarkan hasil dan pembahasan penelitian, maka penitian dapat disimpulkan bahwa citra testing yang berjumlah 20 dan diujikan menghasilkan 19 citra testing mampu dideteksi dengan baik dan 1 citra testing yang tidak dapat dideteksi dengan benar sehingga prosesentase akurasi deteksi tingkat kepadatan rumah penduduk mencapai 95%.

Saran yang dapat diberikan antara lain:

- a. Pengambilan data sebaiknya lakukan saat musim hujan sehingga citra selain rumah akan berwarna hijau

- b. Penelitian ini dapat dikembangkan untuk menghitung objek rumah dan menghitung jarak antar rumah sehingga dapat membandingkan dengan Standar Nasional Indonesia tentang tata cara perencanaan lingkungan perumahan sederhana tidak bersusun dalam rata-rata 50 unit rumah penduduk per hektar.

Yustanti, W., 2012. Algoritma K-Nearest Neighbour untuk Memprediksi Harga Jual Tanah. *Jurnal Matematika statistika dan komputasi*, 9(1), pp.57–68.

DAFTAR PUSTAKA

- Anon 2016. *Kamus Besar Bahasa Indonesia*. [online] Jakarta: Badan Pengembangan dan Pembinaan Bahasa. Available at: <<https://kbbi.kemdikbud.go.id/>>.
- Badan Standar Nasional, 2004. Tata cara perencanaan lingkungan perumahan di perkotaan SNI 03-1733-2004. p.58.
- Kementrian PPN/Bappenas, B.P.S., 2013. *Proyeksi Penduduk Indoensia 2010-2035. Bulletin of Experimental Biology and Medicine*, Jakarta.
- Mahsunah, D., 2013. Analisis Pengaruh Jumlah Penduduk, Pendidikan Dan Pengangguran Terhadap Kemiskinan Di Jawa Timur. *Jurnal Pendidikan Ekonomi (JUPE)*, 1(3), pp.1–17.
- Saputra, W.A. and Arifin, A.Z., 2017. Seeded Region Growing pada Ruang Warna HSI untuk Segmentasi Citra Ikan Tuna. *Jurnal Infotel*, 9(1), p.56.
- Sarie, L.D., Atmaja, R.D., Elektro, F.T., Telkom, U. and Earth, G., 2016. Earth Berbasis Pengolahan Citra Digital Building Count Detection By Using Google Earth Based on Digital. *Earth Berbasis Pengolahan Citra Digital Building Count Detection By Using Google Earth Based on Digital*, 3(1), p.2.
- Surya, R.A., 2016. Ekstraksi Ciri Citra Batik Berdasarkan Tekstur Menggunakan Metode Gray Level Co Occurrence Matrix. *Prosiding, 6 Desember 2016, Vol 2 No. 1*, [online] 2(1), pp.146–150. Available at: <<http://ars.ilkom.unsri.ac.id>>.
- Wijaya, N. and Ridwan, A., 2019. Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors. *Sisfokom*, 08(1), pp.74–78.
- Yousman and Yeyep, 2018. *Google Earth*. Yogyakarta: ANDI.