

PENGELOMPOKKAN *SUNSPOT* PADA CITRA MATAHARI DENGAN MENGGUNAKAN *K-MEANS CLUSTERING*

Rifa Atul Hasanah¹, Dian Candra Rini Novitasari², Nanang Widodo³, Ahmad Hanif Asyhar⁴
Matematika UIN Sunan Ampel Surabaya^{1,2,4}
Lembaga Penerbangan dan Antariksa Nasional Pasuruan³
e-mail : diancrini@uinsby.ac.id

Abstrak – Pada lapisan fotosfer nampak *sunspot* atau bintik matahari yang mana *sunspot* tersebut dapat menimbulkan ledakan-ledakan, seperti ledakan dahsyat (*flare*) dan pelontaran massa korona (*Coronal Mass Ejection/CME*). Ledakan-ledakan ini dapat mengganggu komunikasi radio frekuensi tinggi dan kebisingan radio yang mengganggu komunikasi dan sistem radar. Untuk mengetahui tingkat kompleksitas grup *sunspot* dan aktivitasnya digunakan klasifikasi metode Zurich, yang berisi tentang klasifikasi jenis grup *sunspot*. Informasi ini sangat penting untuk mengetahui seberapa besar gangguan yang didapatkan dari jenis grup *sunspot* tersebut. Tujuan dari penulisan yaitu untuk meneliti bagaimana mengelompokkan *sunspot* pada citra matahari dengan menggunakan *K-Means Clustering*. Pengelompokan *sunspot* menggunakan data citra matahari. Citra matahari diproses untuk diambil posisi x, y . Pengambilan posisi x, y sesuai dengan piksel *sunspot* yang digunakan untuk proses *clustering*. Hasil penelitian yaitu *cluster* piksel *sunspot* yang menunjukkan grup *sunspot*, hasil *clustering* telah menunjukkan hasil yang baik dengan nilai *Silhouette Coefficient* sebesar 0.9381, yang berarti bahwa struktur dari *cluster* termasuk kuat.

Kata Kunci – Matahari, *Sunspot*, *Flare*, Citra

I. PENDAHULUAN

Cuaca antariksa mengacu pada kondisi buruk pada matahari yang dapat memengaruhi sistem teknologi dan dapat membahayakan kesehatan atau kehidupan manusia [1]. Matahari merupakan bintang terdekat dari bumi dan merupakan sumber energi yang tak terbatas bagi kehidupan di

bumi. Aktivitas matahari yang dapat diamati dari bumi adalah aktivitas yang terjadi di lapisan fotosfer, kromosfer, dan korona matahari [2]. Pada lapisan fotosfer ini tidak hanya tampak sinar matahari, namun juga tampak *sunspot* atau bintik matahari. Pada matahari terdapat ledakan-ledakan yang ditimbulkan oleh *sunspot*, ledakan tersebut dapat berupa ledakan dahsyat (*flare*) dan pelontaran massa korona (*Coronal Mass Ejection/CME*) [3]. Aktivitas matahari dapat menghasilkan *flare X-rays* yang dapat mengganggu komunikasi radio frekuensi tinggi dan kebisingan radio yang mengganggu komunikasi dan sistem radar [4]. Kondisi cuaca antariksa yang buruk dapat menyebabkan anomali dan kegagalan sistem dan peningkatan hambatan pada pergerakan satelit dan pesawat ruang angkasa [5].

Bintik matahari adalah daerah yang terlihat di fotosfer Matahari yang mencerminkan aktivitas matahari. Identifikasi *sunspot* beserta sifat-sifatnya seperti ukuran, kontras, dan posisi pada piringan matahari memiliki peran penting dalam memprediksi cuaca ruang angkasa dan untuk mempelajari total radiasi Matahari, rotasi diferensial Matahari, dan pemodelan zona konveksi Matahari, variabilitas dari jari-jari Matahari, dan fenomena penting lainnya [6]. Untuk mengetahui tingkat kompleksitas grup *sunspot* dan aktivitasnya digunakan klasifikasi metode Zurich untuk mengklasifikasikan jenis grup *sunspot*. Informasi ini sangat penting untuk

mengetahui seberapa besar gangguan yang didapatkan dari jenis grup *sunspot* tersebut [7].

Citra digital adalah gambar dua dimensi yang bisa ditampilkan pada layar komputer sebagai himpunan nilai digital yang disebut piksel. Citra mempunyai karakteristik yang tidak dimiliki data teks. Suatu citra dapat didefinisikan sebagai fungsi $f(x,y)$ berukuran M baris dan N kolom, dengan x dan y adalah koordinat (x,y) dinamakan intensitas atau tingkat keabuan dari citra pada titik tersebut [8]. Klasifikasi citra digunakan untuk pengelompokan piksel pada suatu citra kedalam sejumlah kelas, sehingga setiap kelas dapat menggambarkan suatu entitas dengan ciri-ciri tertentu [9]. *Clustering* merupakan proses pengelompokan data dalam kelas-kelas atau *cluster-cluster* sehingga data-data dalam suatu *cluster* memiliki tingkat persamaan yang tinggi satu dengan yang lainnya tetapi sangat berbeda dengan data pada *cluster* lain [10]. *K-Means* merupakan metode *clustering* yang paling sederhana [11] metode ini menggunakan *centroid* untuk membuat *cluster*, *centroid* adalah nilai titik tengah suatu *cluster* dan *centroid* digunakan untuk menghitung jarak suatu objek terhadap *centroid*. Suatu objek data termasuk dalam suatu *cluster* jika memiliki jarak terpendek terhadap *centroid cluster* tersebut [12]. Algoritma *K-Means* ini merupakan salah satu dari algoritma yang banyak digunakan dalam pengelompokan karena kesederhanaan dan efisiensi [13] dan diakui sebagai salah satu dari 10 algoritma data mining teratas oleh IEEE [14].

Pada penelitian yang berjudul Algoritma *K-Means Clustering* Dalam Pengolahan Citra Digital Landsat, pada penelitian itu dilakukan *clustering* untuk mengetahui daerah penutupan lahan, hasil yang diperoleh menunjukkan 94,4% sesuai dengan lapangan dan 5,6% tidak sesuai [15]. *K-Means Clustering* berhasil digunakan dalam penelitian untuk meng-*cluster* provinsi di Indonesia [16]. Waktu untuk proses *clustering* pada algoritma *K-Means* lebih efisien daripada FCM [17]. Pada penelitian kompresi citra MRI, Algoritma *k-means*

dapat bekerja dengan baik pada citra ekstensi “.png”, “.jpg” dan “.bmp” [18]. Sebelumnya pengelompokan *sunspot* pada citra matahari pernah dilakukan dengan menggunakan metode *clustering DBSCAN* dan menghasilkan hasil yang tepat karena hasil pengelompokan *sunspot* sama dengan pengelompokan manual [3]. Berdasarkan latar belakang tersebut, penulis ingin meneliti bagaimana mengelompokkan *sunspot* pada citra matahari dengan menggunakan *K-Means Clustering*.

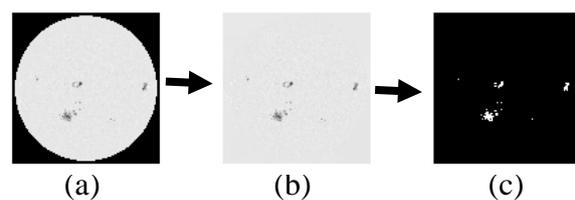
II. HASIL DAN PEMBAHASAN

A. Data

Data yang digunakan dalam penelitian ini menggunakan data citra matahari, yang terdapat *sunspot* dan diambil dari *MDI Continuum / SOHO* untuk proses *clustering* dan data sket grup *sunspot* LAPAN Watukosek untuk membandingkan dengan hasil *clustering*.

B. Pre-Processing

Tahap *pre-processing* pada citra matahari, pinggiran hitam yang berada di luar lingkaran matahari akan dihilangkan karena hanya fokus pada lingkaran matahari yang mana terdapat *sunspot* dengan menggunakan, selanjutnya citra matahari akan diubah menjadi citra biner, pada citra biner warna hitam bernilai 0 dan warna putih bernilai 1, citra biner ini akan digunakan untuk memudahkan dalam proses pengambilan piksel letak *sunspot*. Pada Gambar 1 terlihat bahwa setelah proses binerisasi, maka *sunspot* akan berwarna putih dan mempunyai nilai 1, dan selain *sunspot* akan berwarna hitam dan mempunyai nilai 0, selanjutnya akan diambil data piksel letak *sunspot*, yang akan disimpan dalam bentuk matriks (x,y) dapat dilihat pada Tabel 1.



Gambar 1 : Preprocessing Citra Matahari.
 (a) Citra asli matahari. (b) Citra matahari setelah dihilangkan pinggiran hitamnya. (c) citra biner matahari

Tabel 1 : Contoh posisi x,y piksel *sunspot*

No.	X	Y
1	1296	480
2	1302	558
3	1302	559
4	1303	558

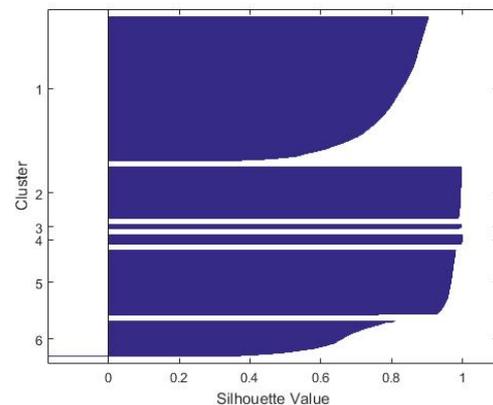
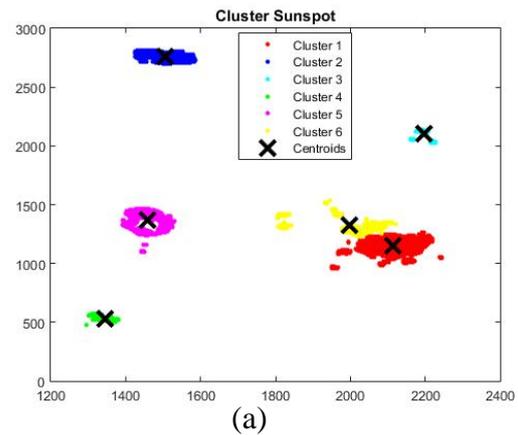
C. Processing

Pada tahap *processing* ini, data piksel letak *sunspot* akan di *clustering* menggunakan *K-Means*. Tujuan dilakukan *clustering* ini adalah untuk menentukan jumlah grup *Sunspot* yang ada dalam satu citra Matahari. Proses *clustering* dengan menggunakan *K-Means* ini dimulai dengan menentukan jumlah *cluster* (k), kemudian menentukan nilai *centroid* secara acak, kemudian menghitung jarak minimum dari data ke *centroid* dengan menggunakan *Euclidean Distance*, setelah itu piksel akan dikelompokkan kedalam *cluster* berdasarkan jarak minimum dengan *centroid*. Selanjutnya, hitung nilai *centroid* baru dengan menggunakan rata-rata dari anggota setiap *cluster*. Jika masih ada anggota dalam suatu *cluster* yang berpindah kedalam *cluster* lain, maka akan dihitung kembali jarak minimum ke *centroid*. Jika sudah tidak ada yang berpindah maka *clustering* selesai. Percobaan *clustering* pada citra Matahari dilakukan pada 15 macam citra matahari. Hasil *clustering* tersebut terlihat pada Tabel 2

Tabel 2 : Hasil *clustering* 15 macam citra matahari

No	Waktu Pengambilan Citra Matahari	Jumlah Cluster
1	08/06/2000	8
2	19/11/2000	8
3	30/04/2003	8
4	02/05/2003	6
5	09/08/2003	4

6	24/09/2003	4
7	21/10/2003	4
8	28/10/2003	6
9	24/12/2003	3
10	16/01/2005	3
11	10/05/2005	5
12	10/06/2005	5
13	12/09/2005	2
14	14/09/2005	1
15	19/11/2005	1

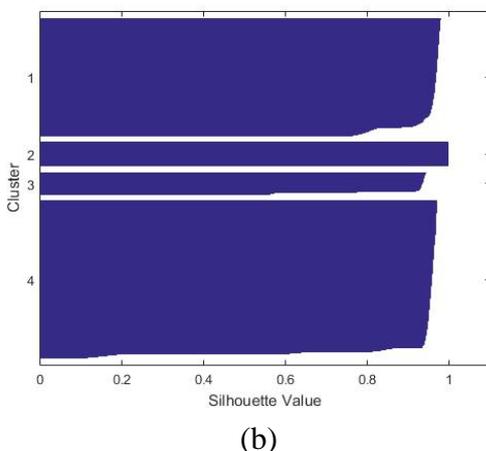
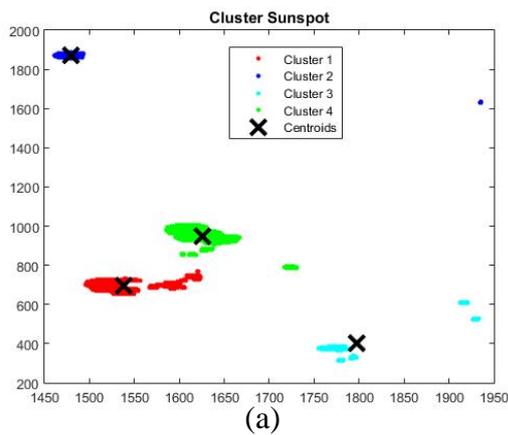


Gambar 2 : (a) Hasil *Clustering* Pada Citra Matahari Pada Tanggal 28/10/2003. (b) Grafik Nilai *Silhouette*

Pada Gambar 2(a) merupakan hasil *clustering* dengan $k = 6$, maka *sunspot* terbagi menjadi 6 kelompok, yaitu *cluster1*, *cluster2*, *cluster3*, *cluster4*, *cluster5* dan *cluster6*.

Setelah proses *clustering* selesai maka akan dilakukan evaluasi *clustering*. Evaluasi *clustering* digunakan untuk mengetahui seberapa baik suatu data terkelompokkan. Pada penelitian ini menggunakan *Silhouette*

Coefficient untuk evaluasi *clustering*. *Silhouette Coefficient* digunakan untuk menemukan kekuatan *cluster* atau mengidentifikasi kualitas *cluster* [19]. Hasil *clustering* dikatakan memiliki struktur *cluster* yang kuat dan berada pada *cluster* yang tepat jika nilai *Silhouette* lebih dari 0 atau positif, dan jika nilai *Silhouette* kurang dari 0 atau minus maka dapat dikatakan memiliki struktur *cluster* yang lemah dan data tersebut tidak berada pada *cluster* yang tepat [20].

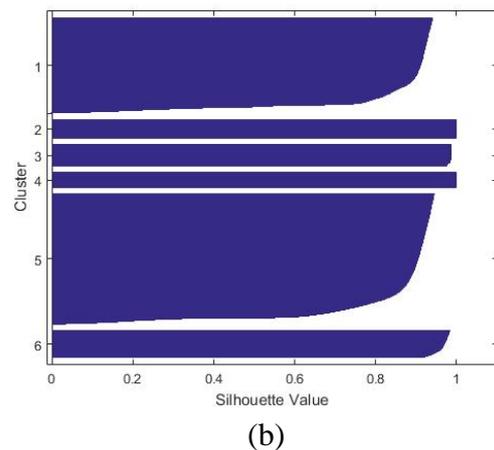
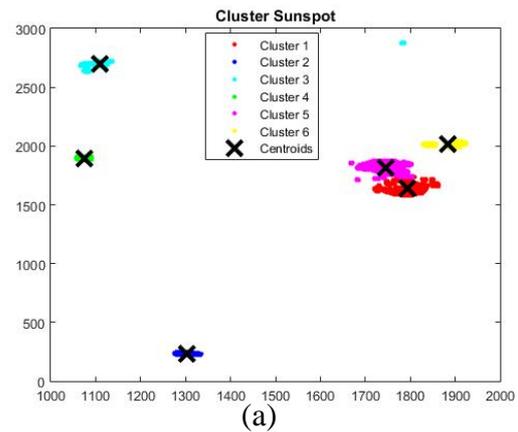


Gambar 3 : (a) Hasil *Clustering* Pada Citra Matahari Pada Tanggal 24/09/2003. (b) Grafik Nilai *Silhouette*

Pada Gambar 2(b) merupakan grafik nilai *Silhouette* dengan $k = 6$. Rata-rata nilai *Silhouette* pada grafik tersebut sebesar 0.8520, yang dapat dikatakan hasil *clustering* memiliki struktur yang kuat. Akan tetapi pada grafik tersebut terdapat data pada *cluster* yang bernilai kurang dari 0 atau minus, yang berarti data tersebut tidak berada pada *cluster* yang tepat.

Pada Gambar 3(a) merupakan hasil *clustering* dengan $k = 4$, maka *sunspot* terbagi menjadi 4 kelompok, yaitu *cluster1*, *cluster2*, *cluster3* dan *cluster4*.

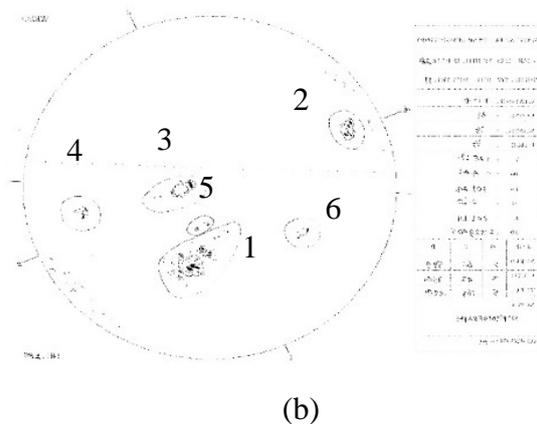
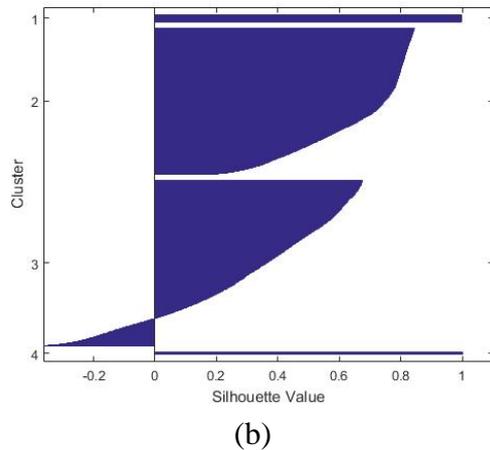
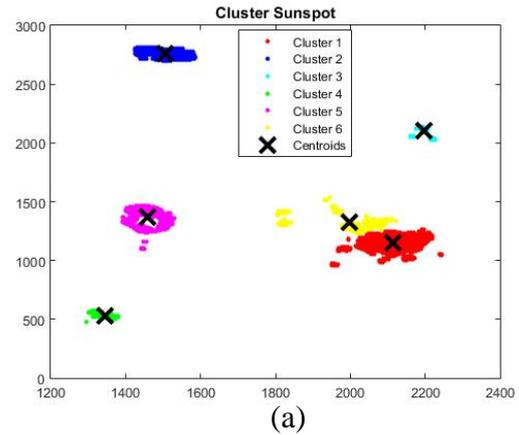
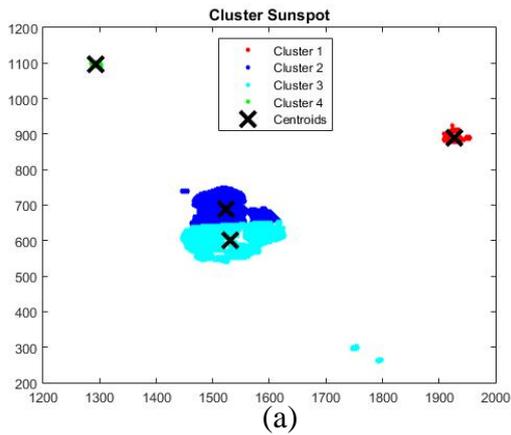
Pada Gambar 3(b) merupakan grafik nilai *Silhouette* dengan $k = 4$. Rata-rata nilai *Silhouette* pada grafik tersebut sebesar 0.9381, yang dapat dikatakan hasil *clustering* memiliki struktur yang kuat.



Gambar 4 : (a) Hasil *Clustering* Pada Citra Matahari Pada Tanggal 02/05/2003. (b) Grafik Nilai *Silhouette*

Pada Gambar 4(a) merupakan hasil *clustering* dengan $k = 6$, maka *sunspot* terbagi menjadi 6 kelompok, yaitu *cluster1*, *cluster2*, *cluster3*, *cluster4*, *cluster5* dan *cluster6*.

Pada Gambar 4(b) merupakan grafik nilai *Silhouette* dengan $k = 6$. Rata-rata nilai *Silhouette* pada grafik tersebut sebesar 0.8825, yang dapat dikatakan hasil *clustering* memiliki struktur yang kuat.



Gambar 5 : (a) Hasil *Clustering* Pada Citra Matahari Pada Tanggal 21/10/2003. (b) Grafik Nilai *Silhouette*

Pada Gambar 5(a) merupakan hasil *clustering* dengan $k = 4$, maka *sunspot* terbagi menjadi 4 kelompok, yaitu *cluster1*, *cluster2*, *cluster3* dan *cluster4*.

Pada Gambar 5(b) merupakan grafik nilai *Silhouette* dengan $k = 4$. Rata-rata nilai *Silhouette* pada grafik tersebut sebesar 0.4984, yang dapat dikatakan hasil *clustering* memiliki struktur yang lemah dan pada grafik tersebut terdapat data pada *cluster* yang bernilai kurang dari 0 atau minus, yang berarti data tersebut tidak berada pada *cluster* yang tepat.

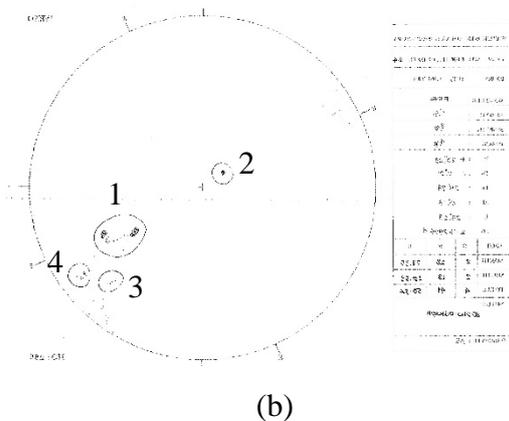
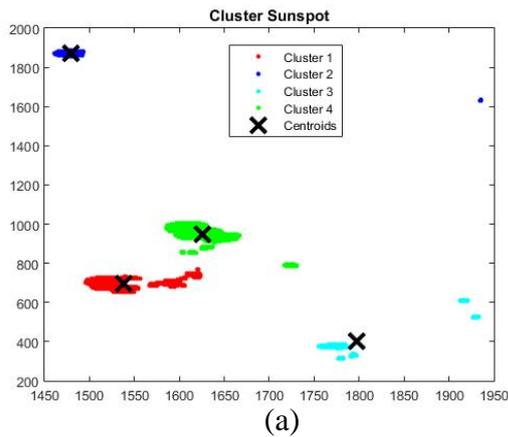
Setelah itu dilakukan perbandingan hasil *clustering* dengan sket *sunspot* LAPAN

Gambar 6 : Perbandingan Hasil *Clustering Sunspot* Pada Citra Matahari Pada Tanggal 28/10/2003 Dengan Sket Grup *Sunspot* LAPAN Pada Tanggal 28/10/2003. (a) *Clustering* Dengan $k = 6$, (b) Sket Grup *Sunspot* LAPAN

Pada sket grup *sunspot* LAPAN terlihat bahwa ada 6 grup *sunspot* yang berlabel 1-6 yang merupakan hasil pengelompokan oleh pakar LAPAN Watukosek.

Pada hasil *clustering* menggunakan *K-Means* dengan $k = 6$ terbagi menjadi 6 grup *sunspot*, yaitu *cluster1*, *cluster2*, *cluster3*, *cluster4*, *cluster5* dan *cluster6*. Hasil *clustering* terlihat bagus karena *cluster2*, *cluster3*, *cluster4* dan *cluster6* mengelompokkan *sunspot* sama seperti pada label 2,3,4, dan 6. Terlihat berbeda pada *cluster1* dan *cluster5* yang mana seharusnya *cluster1* merupakan satu grup *sunspot* yang besar dan *cluster5* merupakan satu grup *sunspot* kecil seperti pada sket grup *sunspot* LAPAN yang berlabel 1 dan 5. Oleh karena itu dapat disimpulkan bahwa, jika dilihat dari banyaknya grup *sunspot* yang terbentuk dari hasil *clustering* maka *clustering* dengan $k=6$

merupakan *clustering* yang paling baik untuk citra *sunspot* pada tanggal 28/10/2003.



Gambar 7 : Perbandingan Hasil *Clustering Sunspot* Pada Citra Matahari Pada Tanggal 24/09/2003 Dengan Sket Grup *Sunspot* LAPAN Pada Tanggal 24/09/2003. (a) *Clustering* Dengan $k = 4$, (b) Sket Grup *Sunspot* LAPAN

Pada sket grup *sunspot* LAPAN terlihat bahwa ada 4 grup *sunspot* yang berlabel 1-4 yang merupakan hasil pengelompokkan oleh pakar LAPAN Watukosek.

Pada hasil *clustering* menggunakan *K-Means* dengan $k = 4$ terbagi menjadi 4 grup *sunspot*, yaitu *cluster1*, *cluster2*, *cluster3* dan *cluster4*. Hasil *clustering* terlihat kurang bagus karena berhasil mengelompokkan satu *sunspot* yang sama seperti hasil pengelompokkan pada sket *sunspot* LAPAN, yaitu *cluster2* sama seperti label 2. Terlihat berbeda pada *cluster1*, *cluster3* dan *cluster4* yang mana seharusnya *cluster1* dan *cluster4* merupakan satu grup *sunspot* yang besar dan *cluster3* merupakan dua grup *sunspot* kecil seperti pada sket grup *sunspot* LAPAN yang berlabel 1, 3 dan 4.

III. KESIMPULAN

Berdasarkan hasil dan pembahasan, maka dapat disimpulkan bahwa pengelompokkan *sunspot* pada citra matahari dengan menggunakan *K-Means Clustering* memiliki hasil yang baik, dengan nilai *Silhouette Coefficient* sebesar 0.9381. Hanya saja ketika terdapat data yang berdekatan, hasil *cluster* tidak sesuai dengan kelas yang sebenarnya. Maka jika terdapat data yang memiliki titik berdekatan diperlukan penentuan titik pusat *cluster* di tahap awal, agar lebih tepat dalam mengelompokkan.

REFERENSI

- [1] T. Colak and R. Qahwaji, "Automated Prediction of Solar Flares Using Neural Networks and Sunspots Associations," pp. 1–9, 2007.
- [2] R. E. Smolinski, *Understanding The Universe*. 2011.
- [3] G. S. Budhi, R. Adipranata, M. Sugiarto, B. Anwar, and B. Setiahadhi, "Pengelompokan Sunspot Pada Citra Digital Mahatari Menggunakan Metode Clustering DBSCAN," in *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 2011.
- [4] M. . Shea and D. . Smart, *Space weather: The effect on operations in space*. in *Solar Terrestrial Relation: Predicting The Effects On The Near-Earth Environment*, 1998.
- [5] F. Guy and D. . Johnson, "Space Environmental Impacts on DoD Operations," *AIR FORCE Sp. Command*, 2003.
- [6] D. Djafer, A. Irbah, and M. Meftah, "Identification of Sunspots on Full-Disk Solar Images Using Wavelet Analysis," *Sol. Phys.*, vol. 281, no. 2, pp. 863–875, 2012.
- [7] R. Adipranata, G. S. Budhi, B. Setiahadhi, and B. Anwar, "Segmentasi Bintik Matahari Menggunakan Watershed," in *Konferensi Nasional Sistem dan Informatika*, 2010, pp. 332–337.
- [8] D. Jatikusumo, *Modul Perkuliahan: Pengolahan Citra*. Universitas Mercu Buana: Pusat Bahan Ajar dan eLearning, 2017.
- [9] A. Z. Arifin and W. D. S. Kurniati, "Penggunaan Analisa Faktor Untuk Klasifikasi Citra Penginderaan Jauh Multispektral," vol. 1, no. 1, 2002.
- [10] D. A. Masatu, I. Soesanti, and H. A. Nugroho, "Penerapan Algoritma Kompresi Jpeg Dan Metode Fuzzy C-Means Pada Kompresi Citra Berbasis," *Penelit. Tek. Elektro dan Teknol. Inf.*, vol. 1, no. 1, pp. 7–11, 2014.

- [11] M. Sammour and Z. Othman, "An Agglomerative Hierarchical Clustering with Various Distance Measurements for Ground Level Ozone Clustering in Putrajaya , Malaysia," vol. 6, no. 6, pp. 1127–1133, 2016.
- [12] B. Simamora, *Analisis Multivariat Pemasaran*. Jakarta: PT. Gramedia Pustaka Utama, 2005.
- [13] J. A. Harding, M. Shahbaz., S. Srinivas, and A. Kusiak, "Data Mining in Manufacturing," 2006.
- [14] X. Wu, "Top 10 Algorithms In Data Mining," 2008.
- [15] N. R. Apriyanti, R. A. Nugroho, and O. Soesanto, "Algortima K-Means Clustering Dalam Pengolahan Citra Digital Landsat," *KLIK Kumpul. J. Ilmu Komput.*, vol. 02, no. 02, pp. 1–13, 2015.
- [16] A. S. Ahmar, D. Napitupulu, and R. Rahim, "Using K-Means Clustering to Cluster Provinces in Indonesia Using K-Means Clustering to Cluster Provinces in Indonesia," 2018.
- [17] A. K. Dubey, U. Gupta, and S. Jain, "Comparative Study of K-means and Fuzzy C-means Algorithms on The Breast Cancer Data," vol. 8, no. 1, pp. 18–29, 2018.
- [18] I. W. A. Wijaya and A. Kusumadewi, "Penerapan Algoritma K-Means Pada Kompresi Adaptif Citra Medis MRI," *INFORMATIKA*, vol. 11, no. 2, 2015.
- [19] V. T. P. Swindiaro, R. Sarno, and D. C. R. Novitasari, "Integration of Fuzzy C-Means Clustering and TOPSIS (FCM-TOPSIS) with Silhouette Analysis for Multi Criteria Parameter Data," *Int. Semin. Appl. Technol. Inf. Commun.*, pp. 463–468, 2018.
- [20] T. Alfina, B. Santosa, and R. Barakbah, "Analisa Perbandingan Metode Hierarchical Clustering , K-means dan Gabungan Keduanya dalam Cluster Data (Studi kasus : Problem Kerja Praktek Jurusan Teknik Industri ITS)," vol. 1, 2012.