

# Deteksi Otomatis Daerah Aktif dari Citra SDO/HMI Magnetogram Matahari Menggunakan Adaptive Centroid Clustering (Automatic Detection of Active Region from SDO/HMI Magnetogram Image Using Adaptive Centroid Clustering)

Ahmad Zulfiana Utama, Tiar Dani

Pusat Sains Antariksa, Bandung

---

## Riwayat naskah

Diterima: 07-12-2020

Direvisi: 16-04-2021

Disetujui: 19-04-2021

## Korespondensi:

Ahmad Zulfiana Utama

ahmad.zulfiana@lapan.go.id

## ABSTRAK

Peristiwa *flare* yang terjadi di daerah aktif Matahari dapat diamati melalui citra yang dihasilkan oleh satelit *Solar Dynamics Observatory* (SDO). Citra *Helioseismic and Magnetic Imager* (HMI) Magnetogram SDO menampilkan informasi area medan magnet pada permukaan Matahari yang direpresentasikan oleh area putih (polaritas positif) dan area hitam (polaritas negatif). Daerah aktif yang muncul dalam citra magnetogram merupakan struktur berpasangan antara polaritas positif dan negatif. Teknik pengolahan citra digital digunakan untuk mendeteksi dan mengelompokkan daerah aktif citra HMI Magnetogram dimulai dari fase *preprocessing*, pengelompokan kluster lokal kontur hitam/putih, pengelompokan daerah aktif, dan ekstraksi parameter luas daerah aktif. Algoritma *Adaptive Centroid Clustering* (ACC) telah dikembangkan untuk mengelompokkan daerah aktif berdasarkan titik tengah (*centroid*) pada kluster yang terdeteksi, jarak antara *centroid*, dan perpotongan garis batas area. Pengujian pengelompokan daerah aktif menggunakan 12 data historis peristiwa *flare* kelas X dan M yang terekam dan diuji menggunakan matriks evaluasi *area-based accuracy*. Matriks evaluasi menunjukkan bahwa algoritma ACC dapat mendeteksi daerah aktif pada citra HMI Magnetogram. Selain itu, setiap daerah aktif yang terdeteksi pada 12 citra dapat diekstraksi parameter luasnya dan diperoleh bahwa luas daerah aktif terbesar memiliki korelasi dengan daerah aktif penghasil *flare*.

**Kata kunci:** Pengelompokan kontur, Pengolahan citra digital, Daerah aktif

## ABSTRACT

Flare events occur in the active regions of the Sun and can be observed by the Solar Dynamics Observatory (SDO) satellite. Images taken by the Helioseismic and Magnetic Imager (HMI) Magnetogram onboard the SDO show the magnetic field on the Sun's surface in which the white (black) regions represent the magnetic positive (negative) polarities. Active region is a paired structure of positive and negative polarity. Active regions are detected and classified by using several image processing techniques, which are preprocessing phase, local cluster grouping of black/white areas, active regions grouping, and extraction of the active region parameters. Adaptive Centroid Clustering (ACC) algorithm has been developed to classify active regions by using area centroids, which are grouped based on the distance and intersection regions. To examine the results, we used 12 historical data of recorded X and M class flare events and tested them by using an area-based accuracy evaluation matrix. The evaluation matrix shows that the ACC algorithm can detect active regions in the HMI Magnetogram. Area of the detected active regions can be extracted from the 12 images. It is found that the largest active region in image is related to the flare productive active region.

**Keywords:** Clustering, Digital image processing, Active region

## 1. Pendahuluan

Aktivitas Matahari berupa *flare* dan *coronal mass ejection* (CME) dapat memberikan dampak terhadap lingkungan antariksa dan lingkungan di sekitar Bumi. Dampak serius akibat peristiwa tersebut di antaranya adalah gangguan terhadap operasional satelit yang berada di orbitnya, gangguan terhadap komunikasi radio HF, gangguan terhadap jaringan listrik, dan juga radiasi, terutama terhadap penerbangan yang melintasi kutub Bumi (Schrijver *et al.*, 2015). Untuk meminimalkan dampak dari aktivitas Matahari tersebut, diperlukan sistem peringatan dini yang andal agar dapat diambil langkah mitigasi untuk melindungi infrastruktur teknologi yang terdampak (Sreeja, 2016).

Daerah aktif memiliki keterkaitan yang erat dengan peristiwa *flare* atau CME. Daerah aktif dapat diamati menggunakan citra *Helioseismic and Magnetic Imager* (HMI) Magnetogram pada satelit *Solar Dynamics Observatory* (SDO) (Bobra *et al.*, 2014). Citra magnetogram tersebut dapat memberikan informasi mengenai kompleksitas medan magnet suatu daerah aktif yang direpresentasikan oleh pasangan polaritas positif dan negatif (Mackay *et al.*, 2011).

Pada contoh kasus yang paling sederhana, daerah aktif terbentuk dari struktur bipolar (positif dan negatif), akan tetapi pada kasus lain daerah aktif dapat berbentuk kompleks dengan berbagai ukuran dan jumlah area magnetik (Guglielmino *et al.*, 2012). Daerah putih dalam citra magnetogram merepresentasikan polaritas positif dan daerah hitam merepresentasikan polaritas negatif, sedangkan daerah berwarna abu-abu mengindikasikan daerah yang memiliki kekuatan medan magnet yang lemah.

Hubungan kompleksitas daerah aktif dengan produktifitas *flare* telah diteliti oleh Georgoulis (2008). Beliau mempelajari peristiwa erupitif pada 23 daerah aktif dan menemukan keterkaitan fitur batas garis antara daerah polaritas positif dan negatif, atau dikenal dengan *polarity inversion line* (PIL), dengan kemunculan *flare* atau CME.

Steward *et al.* (2017) mengembangkan program Flarecast untuk mendeteksi dan mengambil fitur daerah aktif menggunakan citra magnetogram *Global Oscillation Network Group* (GONG) untuk memprediksi peristiwa *flare* menggunakan model probabilitas pembelajaran mesin. Data garis-pandang GONG digunakan untuk mendeteksi daerah

aktif berdasarkan nilai ambang tertentu. Selanjutnya, dicari nilai PIL pada daerah yang terdeteksi menggunakan teknik *shifting-pixel*. PIL dan fitur daerah aktif digunakan sebagai masukan pada model probabilitas sebagai data latih dalam memprediksi peristiwa *flare*.

Aplikasi SMART (*Solar Feature Detection Algorithms*) mengekstraksi karakteristik daerah aktif untuk memahami keterkaitannya dengan kemunculan peristiwa *flare* (Higgins *et al.*, 2011). Aplikasi ini membandingkan dua buah citra SOHO/MDI LOS Magnetogram dengan interval waktu 96 menit. Pengaplikasian teknik pengolahan citra seperti *smoothing*, *noise-threshold*, *static-threshold*, *line of sight correction*, *masking*, *subtraction*, dan *dilation* dilakukan secara berulang terhadap dua citra tersebut untuk memperoleh *indexed grown mask* (IGM) dari suatu daerah aktif. IGM merupakan kumpulan nilai piksel yang membentuk kontur pada daerah aktif yang terdeteksi. Ekstraksi fitur daerah aktif diperoleh berdasarkan nilai IGM. Hasil akhir dari aplikasi ini menghasilkan katalog struktur data daerah aktif yang terdeteksi beserta fitur yang terkandung di dalamnya.

Algoritma Ashamari *et al.* (2015) menggunakan citra intensitigram dan magnetogram MDI-SOHO dalam mendeteksi suatu daerah aktif. Pemrosesan dimulai dari deteksi awal bintik Matahari pada citra intensitigram dan deteksi daerah magnetik pada citra magnetogram. Metode *thresholding* dan *mean filter* digunakan untuk mendeteksi daerah bintik Matahari. Sementara itu, daerah magnetik diperoleh dari proses *contrast stretching*, *median-filter*, dan *thresholding*. Nilai batas pada proses *thresholding* diperoleh dari analisis data statistik pada periode minimum aktivitas Matahari. Daerah aktif diperoleh berdasarkan lokasi awal bintik Matahari yang terdeteksi, lalu dibandingkan dengan lokasi daerah magnetik. Jika lokasi bintik Matahari berada di lokasi yang sama maka daerah magnetik dijadikan sebagai kandidat daerah aktif. Atribut kandidat daerah aktif dan bintik Matahari dimasukkan ke dalam metode *neural network* (NN) sebagai data latihan. Keputusan akhir dalam menentukan daerah aktif ditentukan oleh hasil metode NN. Tujuan utama dari algoritma deteksi dan ekstraksi fitur daerah aktif dan daerah magnetik yang sudah dilakukan oleh Ashamari *et al.* (2015), Higgins *et al.* (2011), dan Georgoulis (2018) adalah untuk mengetahui karakteristik dan memahami

keterkaitan peristiwa *flare* berdasarkan karakteristik daerah aktif.

Daerah aktif dapat muncul lebih dari satu atau tidak sama sekali dalam satu citra HMI Magnetogram. Program harus mampu untuk membedakan, memberikan tanda, dan mengelompokkan tiap daerah aktif. Penelitian ini akan menjabarkan deteksi daerah aktif pada citra HMI Magnetogram menggunakan teknik pengolahan citra dan penggunaan sebuah algoritma baru yang diberi nama *Adaptive Centroid Clustering (ACC)*. Algoritma ACC ini berfungsi untuk mengelompokkan daerah aktif berdasarkan informasi *centroid* dari kontur polaritas negatif dan positif yang terdeteksi. Pengelompokan tersebut dilakukan berdasarkan jarak *centroid*, garis perpotongan luas daerah, penggabungan daerah, dan optimasi pengelompokan kluster secara berulang.

Proses identifikasi daerah aktif pada penelitian ini relatif lebih singkat jika dibandingkan dengan penelitian Higgins *et al.* (2011) dan Ashamari *et al.* (2015). Hal tersebut dikarenakan pada penelitian mereka pencarian daerah aktif menggunakan dua buah tipe citra yang berbeda. Selain itu, penggunaan fungsi pencarian dan penggabungan kontur pada pustaka pemrograman OpenCV dapat mempercepat proses identifikasi daerah aktif. Fungsi kontur dikombinasikan dengan algoritma ACC dalam mendeteksi dan mengelompokkan daerah aktif berdasarkan daerah polaritas negatif dan positif pada citra HMI Magnetogram.

Verifikasi hasil deteksi daerah aktif menggunakan metode tingkat akurasi deteksi daerah yang digunakan pada penelitian *object-based image analysis* (Cai *et al.*, 2018). Selanjutnya, parameter luas daerah aktif yang terdeteksi akan diekstraksi dan dikorelasikan terhadap 12 peristiwa *flare*.

## 2. Metodologi

### 2.1. Data

Penelitian ini menggunakan citra magnetogram *full-disk* yang direkam oleh instrumen HMI pada satelit SDO. Citra HMI Magnetogram diperoleh dari website resmi SDO dengan resolusi citra 512 × 512 piksel.

Data pengujian menggunakan 12 citra HMI Magnetogram SDO pada saat peristiwa *flare* kelas X dan M selama siklus Matahari ke-24. Hasil deteksi daerah aktif dan luas area akan dikorelasikan dengan 12 data tersebut.

Citra magnetogram memiliki informasi mengenai area medan magnet (daerah aktif)

yang direpresentasikan oleh pasangan polaritas positif dan negatif. Parameter yang akan diambil dari informasi area medan magnet yaitu luas area polaritas negatif (area hitam) dan positif (area putih), luas area total, dan posisi koordinat dari area yang berpasangan (negatif dan positif).

### 2.2. Metode Penelitian

Metode pengelompokan (kluster) sederhana, contohnya algoritma K-Means, tidak dapat mencapai hasil secara maksimal karena kompleksitas objek daerah aktif yang terdiri dari pasangan area positif dan negatif.

Penggunaan algoritma K-Means harus menentukan jumlah kluster yang diinginkan sehingga hasil kluster bergantung pada jumlah inisialisasi tersebut (Kaushik & Mathur, 2014). Menurut Kaushik & Mathur (2014), apabila jumlah kluster yang ditentukan tidak tepat maka hasil yang diperoleh tidak akan maksimal. Penentuan jumlah kluster tersebut tidak cocok digunakan untuk mencari daerah aktif pada citra HMI Magnetogram secara otomatis karena citra tersebut merupakan data observasi yang bersifat dinamis. Selain pengelompokan daerah aktif, program mampu mengekstraksi informasi luas daerah aktif dan koordinat Kartesian berdasarkan piksel citra yang teramati.

Algoritma ACC yang dikembangkan pada penelitian ini tidak perlu menginisialisasi jumlah kluster seperti pada algoritma K-Means, akan tetapi jumlah kluster dicari berdasarkan nilai *centroid* yang terdeteksi secara dinamis.

Algoritma ACC terdiri dari beberapa fase yaitu *preprocessing*, pengelompokan kluster lokal kontur area hitam dan putih, pengelompokan daerah aktif, dan ekstraksi parameter. Fase *preprocessing* bertujuan untuk mendapatkan informasi piksel dari setiap kontur area hitam dan putih dengan cara membaca masukan citra HMI Magnetogram dan mengonversikan ke dalam kanal warna yang disimpan dalam larik, konversi citra berwarna ke dalam citra abu-abu dan *binary* (citra hitam putih), menghilangkan derau (*noise*) pada citra, klasifikasi area hitam dan putih berdasarkan nilai batas (*threshold,  $T_p$* ), eliminasi pinggiran Matahari, dan mendapatkan nilai kontur dan *centroid*. Seluruh fungsi tersebut disediakan oleh pustaka openCV pada bahasa pemrograman python (Howse, 2013). Tahapan fase *preprocessing* diilustrasikan pada diagram alir Gambar 1.

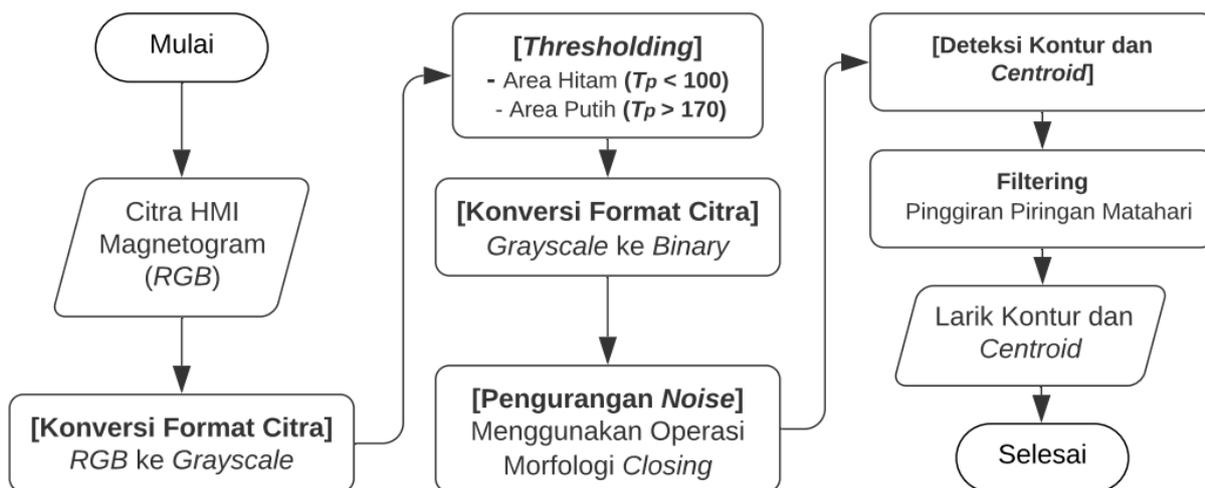
Konversi citra mengubah 3 kanal citra berwarna menjadi 1 kanal warna keabuan (Hidayatullah, 2017). Satu kanal warna menyimpan 8-bit data sehingga konversi citra keabuan akan menghemat ruang penyimpanan data, mengurangi kompleksitas perhitungan, mempermudah pencarian nilai batas untuk klasifikasi area hitam/putih, dan mempercepat perhitungan piksel. Derajat keabuan citra memiliki rentang nilai 0-255 yang didapat dari nilai rata-rata kanal citra berwarna RGB (*red, green, blue*). Piksel dengan nilai terendah akan berwarna hitam, sedangkan nilai tertinggi berwarna putih.

Terdapat derau yang terdistribusi pada citra HMI Magnetogram dan dapat memengaruhi proses pengelompokan daerah aktif. Derau dihilangkan menggunakan operasi morfologi *closing* dan *median filter* dengan setiap piksel dikonvolusi menggunakan sebuah filter  $3 \times 3$  dan nilai piksel yang terkonvolusi digantikan oleh nilai tengah berdasarkan nilai-nilai piksel tetangga (Hidayatullah, 2017).

Klasifikasi area hitam dan putih dilakukan berdasarkan nilai batas,  $T_p$ , pada kanal citra keabuan. Nilai  $T_p$  diambil berdasarkan nilai puncak histogram pada fase Matahari tenang, seperti yang dilakukan oleh Higgins *et al.* (2011), karena nilai puncak tersebut merupakan daerah keabuan yang sebagian besar menyelimuti piringan

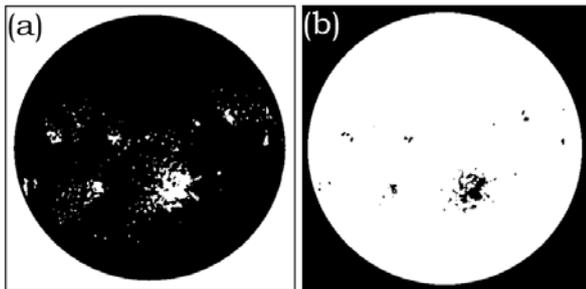
Matahari. Nilai puncak histogram pada fase Matahari tenang adalah 125-130. Nilai tersebut dijadikan acuan sebagai nilai batas untuk membedakan area hitam, putih, dan daerah keabuan. Apabila nilai  $T_p$  lebih dari rentang 125-130, maka daerah keabuan di sekitarnya ikut terdeteksi. Sementara itu, untuk nilai  $T_p$  kurang dari 125-130 terdeteksi sebagai daerah hitam. Penggunaan  $T_p = 120$  menghasilkan daerah hitam beserta derau, sedangkan penggunaan  $T_p = 50$  menghasilkan daerah hitam dengan ukuran yang lebih kecil dan memiliki perbedaan morfologi jika dibandingkan dengan citra awal. Atas pertimbangan tersebut, penelitian ini menggunakan  $T_p = 100$  untuk mendeteksi daerah hitam, karena dapat mempertahankan bentuk morfologi dan meminimalkan derau yang terdeteksi. Persaman (1) menjabarkan bahwa nilai piksel lebih dari  $T_p$  akan dimaksimalkan menjadi 255, dan sebaliknya, area lain diberi nilai 0. Hasil dari proses ini menghasilkan citra *binary* dimana nilai 255 dikonversi menjadi 1 yang menggambarkan piksel berwarna putih dan nilai 0 merepresentasikan piksel berwarna hitam.

$$D_{(x,y)} = \begin{cases} 255, & \text{if } D_{(x,y)} > T_p ; \\ 0, & \text{otherwise} \end{cases} \quad (1)$$



Gambar 1. Diagram alir pada fase *preprocessing* untuk mengklasifikasikan area hitam/putih, dan mencari nilai kontur dan *centroid* pada setiap area yang terdeteksi.

Pencarian area putih mengulangi proses tersebut dengan kondisi yang terbalik, yaitu nilai piksel kurang dari nilai batas akan diberi nilai 0, sedangkan area lainnya diberi nilai 255. Selain itu pencarian area putih menggunakan nilai batas  $T_p = 170$  yang diperoleh berdasarkan nilai histogram pada fase tenang. Pada proses klasifikasi ini akan dihasilkan dua buah citra area hitam dan area putih seperti yang terlihat pada Gambar 2.



Gambar 2. Hasil dari proses klasifikasi area hitam dan putih dalam kanal citra *binary* pada fase *preprocessing*.

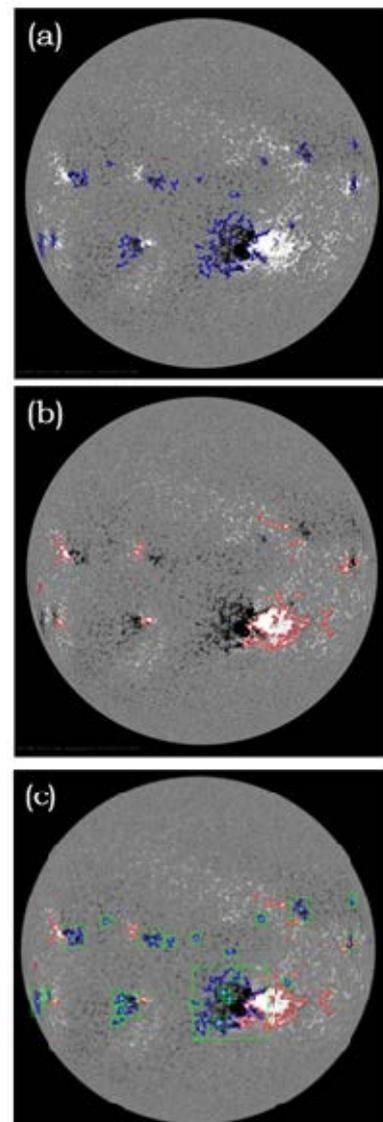
Citra hasil klasifikasi hitam dan putih diproses lebih lanjut untuk mendapatkan informasi kontur dari setiap area magnetik di piringan citra Matahari. Kontur area pada citra merupakan kumpulan piksel yang terkoneksi dan memiliki nilai intensitas warna yang berbeda dengan area tetangga. Pencarian informasi kontur menggunakan fungsi *findContours()* dari pustaka OpenCV. Fungsi *findContours()* menggunakan algoritma yang dikembangkan oleh Suzuki (1985) yang menjadi cikal bakal dalam menganalisis hubungan hierarki dari garis tepi. Hasil kontur area adalah himpunan nilai koordinat  $(x,y)$  garis tepi sebuah objek yang disimpan ke dalam sebuah larik kontur. Gambar 3 merupakan hasil deteksi kontur pada area hitam (a) dan putih (b).

Luas kontur yang terdeteksi dihitung menggunakan fungsi *cv.contourArea()*. Luas kontur digunakan juga untuk mengeliminasi piringan Matahari yang memiliki luas sebesar 174614,5. Nilai ini juga diperoleh dari fungsi *cv.contourArea()*. Nilai *centroid* untuk setiap area yang terdeteksi merupakan nilai tengah dari garis batas kontur area. Kotak hijau pada Gambar 3(c) merupakan garis batas pada setiap kontur dan titik merah merupakan nilai *centroid*. Garis batas memiliki bentuk persegi yang mengelilingi setiap kontur dan digambar menggunakan fungsi *cv.rectangle()*. Untuk mendapatkan titik piksel kiri-atas batas kontur

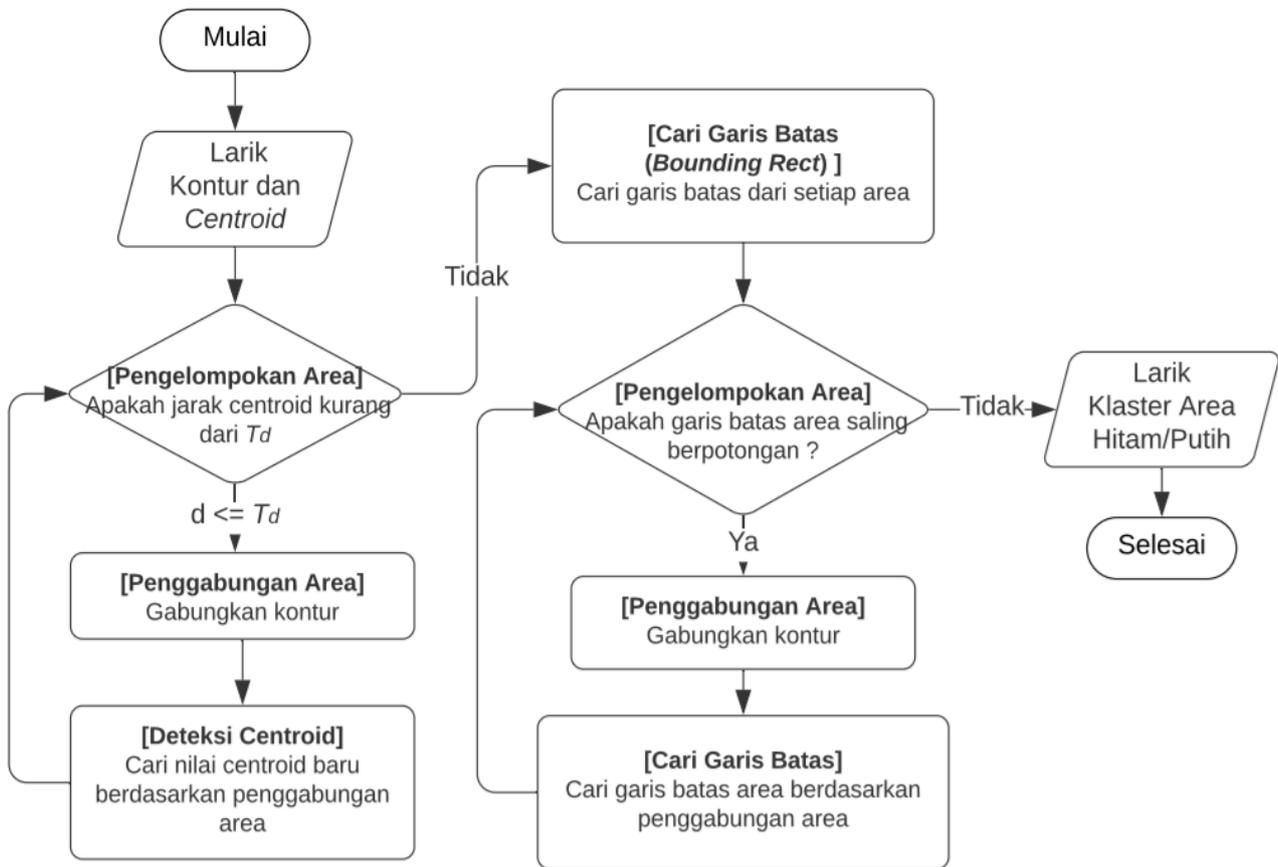
$(x_{min}, y_{min})$ , lebar dan tinggi adalah nilai piksel yang terdapat pada larik kontur menggunakan fungsi *cv.boundingRect()*. Sementara itu, untuk nilai *centroid* diperoleh menggunakan Persamaan (2).

$$Centroid = \begin{cases} Cx = x + 0.5 * width, \\ Cy = y + 0.5 * height \end{cases} \quad (2)$$

Nilai *centroid* dan garis batas kontur area hitam dan putih digunakan untuk mengelompokkan area sejenis (klaster lokal) yang saling bertetangga. Area yang terdeteksi dikelompokkan sesuai dengan jenisnya (hitam/putih) berdasarkan jarak *centroid* dan algoritma perpotongan garis batas kontur.



Gambar 3. (a) Hasil deteksi kontur area hitam (garis biru) (b) Kontur area putih (garis merah) (c) Penggabungan kontur hitam dan putih dalam satu citra.



Gambar 4. Diagram alir pengelompokan kluster lokal pada area hitam dan putih.

Gambar 4 merupakan tahapan dalam mengelompokkan kluster lokal pada area hitam dan putih. Algoritma kluster lokal mengadopsi algoritma K-Means Clustering, yaitu pengelompokan kluster berdasarkan jarak antar nilai *centroid*. Jarak *centroid*,  $d$ , dihitung menggunakan rumus *Euclidian Distance*, seperti pada Persamaan (3).

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3)$$

Nilai *centroid* pada setiap kontur dihitung jaraknya dengan nilai *centroid* lainnya. Pada penelitian ini kami menggunakan nilai batas jarak,  $T_d$ , sebesar 20. Jika  $d$  kurang dari  $T_d$  maka kedua area tersebut digabungkan sehingga akan menghasilkan nilai *centroid* yang baru. Proses penggabungan larik kontur menggunakan fungsi Numpy *concat()*. Proses ini dilakukan secara berulang dan berhenti pada saat jarak antara *centroid* tidak ada yang kurang dari  $T_d$ .

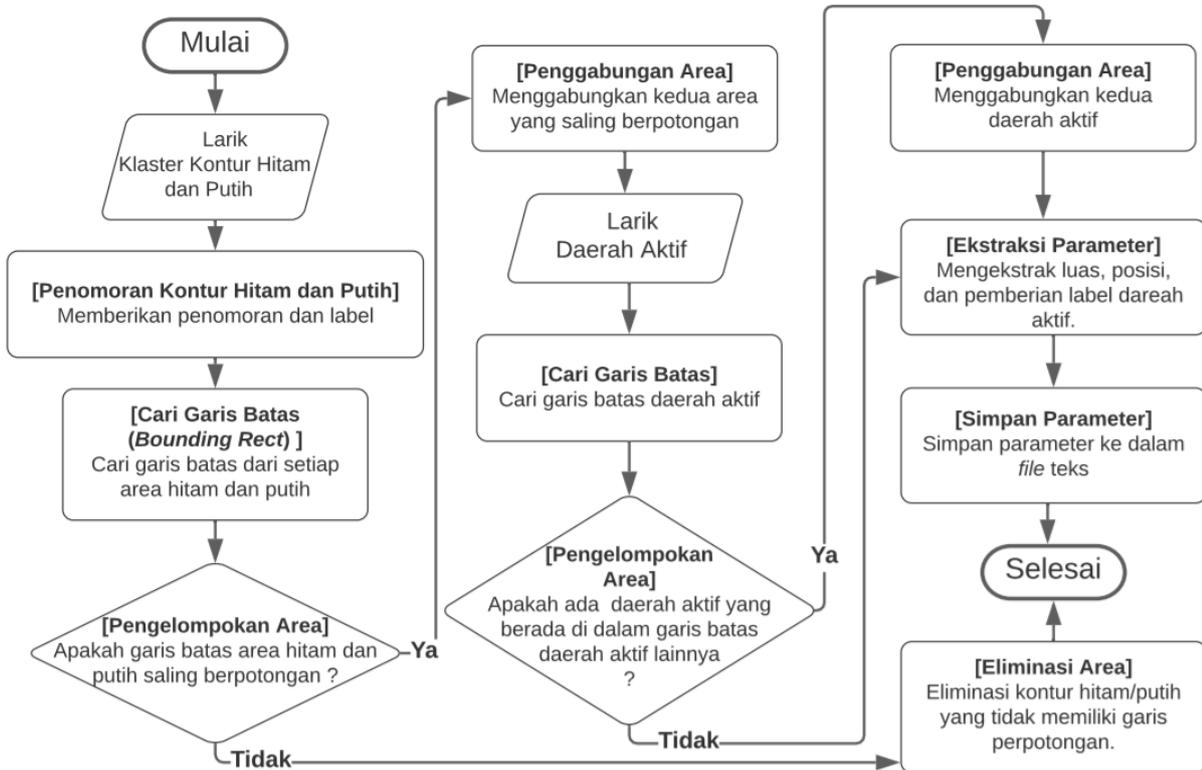
Garis batas kontur yang saling berpotongan merepresentasikan area yang saling bertetangga dan akan dikelompokkan ke dalam satu kluster. Algoritma *intersection region* mendeteksi perpotongan garis dengan

membandingkan nilai koordinat minimum/titik piksel pada kiri-atas ( $x_{min}, y_{min}$ ), dan maksimum/titik piksel kanan-bawah garis batas kontur area ( $x_{max}, y_{max}$ ). Algoritma ini akan dijalankan setelah pengelompokan area berdasarkan jarak *centroid* selesai.

**Algoritma 1. Intersection Region**

```
Function is_intersect (P1, P2)
    if P1.xmin > P2.xmin or P1.xmax < P2.xmax :
        return False
    if P1.ymin > P2.ymin or P1.ymax < P2.ymax :
        return False
    return True
```

Pencarian kluster lokal menghasilkan kluster kontur area hitam dan putih. Informasi ini digunakan dalam pencarian daerah aktif yang terdiri dari pasangan polaritas negatif (area hitam) dan positif (area putih). Gambar 5 merupakan tahapan pengelompokan daerah aktif yang dimulai dari pemberian label unik berdasarkan jenis area untuk setiap kontur untuk menghindari pengelompokan kontur dengan polaritas yang sama. Pemberian label juga memperjelas identitas tiap kontur yang mendefinisikan jenis dan nomor urutan kontur yang terdeteksi.



Gambar 5. Diagram alir pengelompokan daerah aktif berdasarkan kontur hitam dan putih.

Perpotongan garis batas antara kontur area hitam dan putih dijadikan dasar untuk mengelompokkan daerah aktif. Jika garis batas area hitam dan putih berpotongan maka kedua larik kontur area tersebut digabungkan menggunakan fungsi Numpy *concat()*. Penggabungan kontur area hitam dan putih akan menghasilkan garis batas area yang baru yang merepresentasikan suatu daerah aktif. Selanjutnya dilakukan pengecekan apakah ada daerah aktif yang berada di dalam garis batas daerah aktif lainnya. Jika ada, kedua informasi daerah aktif tersebut digabungkan. Proses terakhir adalah ekstraksi dan menyimpan parameter ke dalam fail teks.

Ekstraksi parameter, seperti luas dan posisi daerah aktif dalam koordinat Kartesian, diperoleh dari daerah aktif yang berhasil teridentifikasi. Setiap daerah aktif yang terdeteksi diberikan suatu identitas berupa nomor untuk memudahkan peneliti atau pengguna dalam mencerna informasi.

Luas kontur area daerah aktif dihitung menggunakan fungsi *cv.contourArea()* yang dimiliki oleh OpenCV Python, akan tetapi fungsi tersebut masih dalam satuan luas piksel. Persamaan (4) dapat mengonversikan luas tersebut menjadi satuan *Solar Millionth Hemisphere* (SMH) yang sudah termasuk koreksi *foreshortening* (Smart, 1971; Duffet-

Smith P, 1988; Meadows, 2020). Persamaan (4) digunakan untuk menghitung luas area grup bintang Matahari dengan faktor koreksi *foreshortening*.

$$A_M = \frac{A_s 10^6}{2\pi R^2(\rho)} \quad (4)$$

dengan  $A_M$  adalah luas dalam SMH,  $\rho$  adalah jarak sudut dari permukaan Matahari di tengah piringan ke grup bintang yang diperoleh dari Persamaan (5),  $A_s$  adalah area bintang Matahari yang terukur dalam piksel<sup>2</sup>, dan  $R$  adalah radius Matahari pada citra dalam piksel, dengan  $x$  dan  $y$  merupakan posisi *centroid* dari grup bintang dengan koordinat (0,0) berada di tengah piringan Matahari. Koordinat *default* (0,0) pada citra selalu berada di pojok kiri atas, tetapi untuk perhitungan luas area, kami menggeser koordinat (0,0) tersebut tepat berada di tengah piringan Matahari sesuai dengan syarat Persamaan (6). *appdiam* merupakan diameter semu Matahari (0,53°) dan *diskdiam* adalah diameter Matahari dalam citra (satuan sama dengan  $x$  dan  $y$ ).

$$\rho = (2\rho_1 / \text{appdiam}) - \rho_1 \quad (5)$$

$$\rho_1 = \frac{\text{appdiam}}{\text{diskdiam}} \sqrt{(x^2 + y^2)} \quad (6)$$

Seluruh parameter yang telah berhasil diekstraksi, disimpan ke dalam fail teks. Fail teks berisi sebuah tabel yang terdiri dari 6 kolom, yaitu penomoran daerah aktif, posisi *centroid* dalam arah horizontal  $C_x$ , posisi *centroid* dalam arah vertikal  $C_y$ , luas polaritas negatif (area hitam), luas polaritas positif (area putih), dan total area. Posisi *centroid* daerah aktif dalam nilai koordinat Kartesian dinormalisasi ke dalam sistem koordinat Matahari, dikenal sebagai *heliocentric cartesian position*  $(x,y)$ . Normalisasi tersebut berfungsi untuk memindahkan titik pusat koordinat  $(0,0)$  di pojok kiri atas ke titik tengah piringan Matahari.

**2.3. Metode Pengujian**

Matriks evaluasi untuk akurasi program dalam mendeteksi daerah aktif menggunakan persentase *correctness*, *completeness*, dan *quality*. Metode evaluasi ini telah digunakan untuk mengukur tingkat akurasi deteksi area pada aplikasi *Object-Based Image Analysis* (Cai et al., 2018).

*Correctness*,  $P_{AC}$ , merupakan rasio dari daerah aktif yang terdeteksi sesuai dengan rujukan data historis yang tercatat pada web <http://solarmonitor.org> yang dinotasikan sebagai  $AC$  dan keseluruhan daerah aktif yang terdeteksi oleh program ( $ADC$ ). Evaluasi *correctness* mengukur akurasi program dalam mendeteksi daerah aktif.

$$P_{AC} = \frac{AC}{ADC} \tag{7}$$

*Completeness*,  $P_{AR}$ , merupakan perbandingan  $AC$  dengan keseluruhan daerah aktif  $ARC$  yang ada pada tabel jawaban atau data historis. Tujuannya adalah untuk mengevaluasi kelengkapan daerah aktif yang terdeteksi oleh program.

$$P_{AR} = \frac{AC}{ARC} \tag{8}$$

*Quality*,  $P_{AL}$ , merupakan nilai tengah dari *correctness* dan *completeness* untuk mengukur kualitas hasil deteksi daerah aktif. Pada Persamaan (7), jika nilai  $ADC$  besar dapat berakibat pada kecilnya nilai *correctness*, sedangkan pada persamaan (8), jika nilai  $ARC$  kecil akan menghasilkan *completeness* yang tinggi. Oleh karena itu, persamaan (9) digunakan untuk menyeimbangkan nilai  $P_{AC}$  dan  $P_{AR}$ .

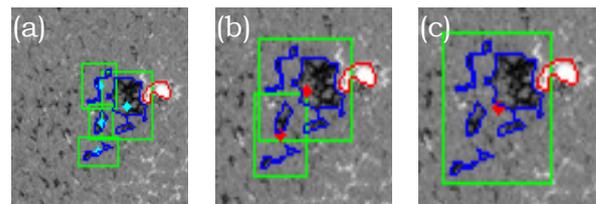
$$P_{AL} = \frac{AC}{ADC + ARC - AC} \tag{9}$$

Pada penelitian ini kami mengeliminasi nilai pada tabel jawaban jika daerah aktif yang terdeteksi lokasinya berada di pinggiran piringan Matahari.

**3. Hasil Pembahasan**

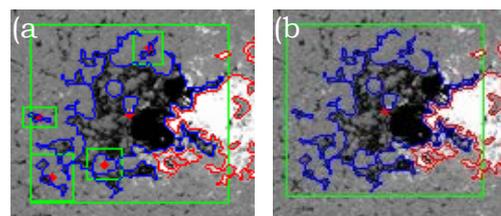
**3.1. Pencarian Kluster Lokal pada Area Hitam dan Putih**

Gambar 6 merupakan ilustrasi dari proses iterasi berdasarkan jarak *centroid*. Pada setiap iterasi, apabila jarak *centroid* kurang dari nilai batas maka kedua kontur tersebut akan dimasukkan ke dalam satu kluster. Kontur area keduanya digabungkan sehingga akan menghasilkan nilai *centroid* yang baru karena perubahan luas kontur. Proses ini dilakukan secara berulang sampai tidak ada jarak *centroid* yang kurang dari nilai batas  $T_d$ .

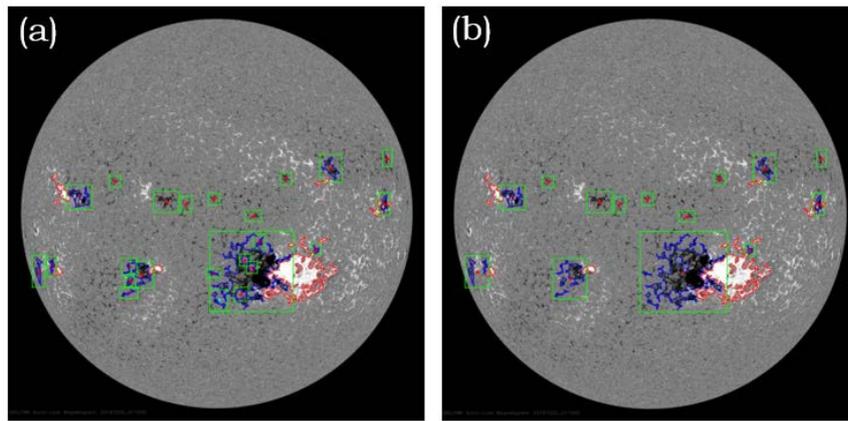


Gambar 6. (a) Iterasi ke-1, (b) iterasi ke-2 sampai (c) iterasi ke- $n$  menghasilkan 1 kluster. Proses berhenti apabila  $d$  lebih dari  $T_d$ .

Setelah dilakukan pengelompokan berdasarkan jarak *centroid*, masih terdapat penumpukan kontur yang berada pada kontur lainnya, seperti ditunjukkan oleh Gambar 7(a). Hal tersebut dikarenakan nilai  $d$  melebihi nilai batas. Oleh karena itu, algoritma perpotongan antara garis batas area akan mengatasi permasalahan tersebut, seperti yang diilustrasikan pada Gambar 7(a) dan (b).



Gambar 7. Cara kerja algoritma *intersection region*. (a) Keadaan awal dan (b) sesudah dilakukan pengelompokan menggunakan algoritma *intersection region*.



Gambar 8. (a) Keadaan awal dan (b) sesudah pengelompokan kluster lokal pada area hitam. Kotak hijau pada panel (a) dan (b) merupakan garis batas area pada setiap kontur dan titik merah merupakan *centroid* dari tiap kontur.

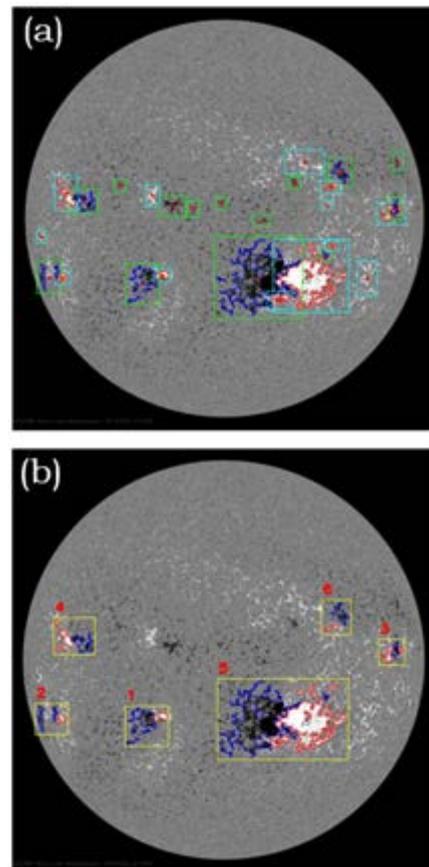
Untuk mengetahui garis batas kontur yang saling berpotongan, dicari nilai koordinat minimum ( $x_{min}, y_{min}$ ) dan maksimum ( $x_{max}, y_{max}$ ) dari setiap batas kontur. Selanjutnya, untuk setiap garis batas yang saling berpotongan, dilakukan penggabungan larik kontur kembali menggunakan fungsi Numpy *concat()*. Penggabungan kontur ini akan menghasilkan garis batas area yang baru sehingga proses algoritma berdasarkan perpotongan garis dilakukan secara berulang seperti yang ditunjukkan diagram alir pada Gambar 5. Seluruh tahapan pada pengelompokan kluster kontur area hitam dilakukan juga pada pengelompokan kontur area putih.

Gambar 8(a) merupakan hasil deteksi kontur area hitam sebelum dikelompokkan berdasarkan jarak terdekat dari *centroid* dan perpotongan garis batas area. Sementara itu, Gambar 8(b) merupakan hasil dari pengelompokan kluster area hitam dan kontur yang tervisualisasi lebih sedikit dibandingkan dengan Gambar 8(a) karena telah dilakukan penggabungan kontur area apabila area hitam/putih berada dalam kluster yang sama.

### 3.2. Pengelompokan Daerah Aktif

Pencarian daerah aktif dilakukan dengan memanfaatkan informasi dari kluster kontur area hitam dan putih yang berpasangan. Penentuan pasangan tersebut berdasarkan perpotongan garis batas kluster area hitam dan putih. Gambar 9(b) adalah hasil proses iterasi ke-*n* pada pencarian daerah aktif yang ditandai oleh garis batas kontur berwarna kuning. Pada setiap proses, apabila ada garis batas kontur area putih dan hitam yang berpotongan maka kedua kontur tersebut digabungkan ke dalam satu kluster

yang merepresentasikan satu daerah aktif. Proses ini dilakukan secara berulang sampai tidak ada garis yang saling berpotongan. Gambar 9(b) adalah citra akhir setelah proses eliminasi kontur hitam dan putih yang tidak termasuk ke dalam daerah aktif atau tidak memiliki pasangan.



Gambar 9. (a) Kluster kontur lokal sebagai masukan pada algoritma pencarian daerah aktif. (b) Iterasi terakhir pengelompokan daerah aktif berdasarkan kontur lokal.

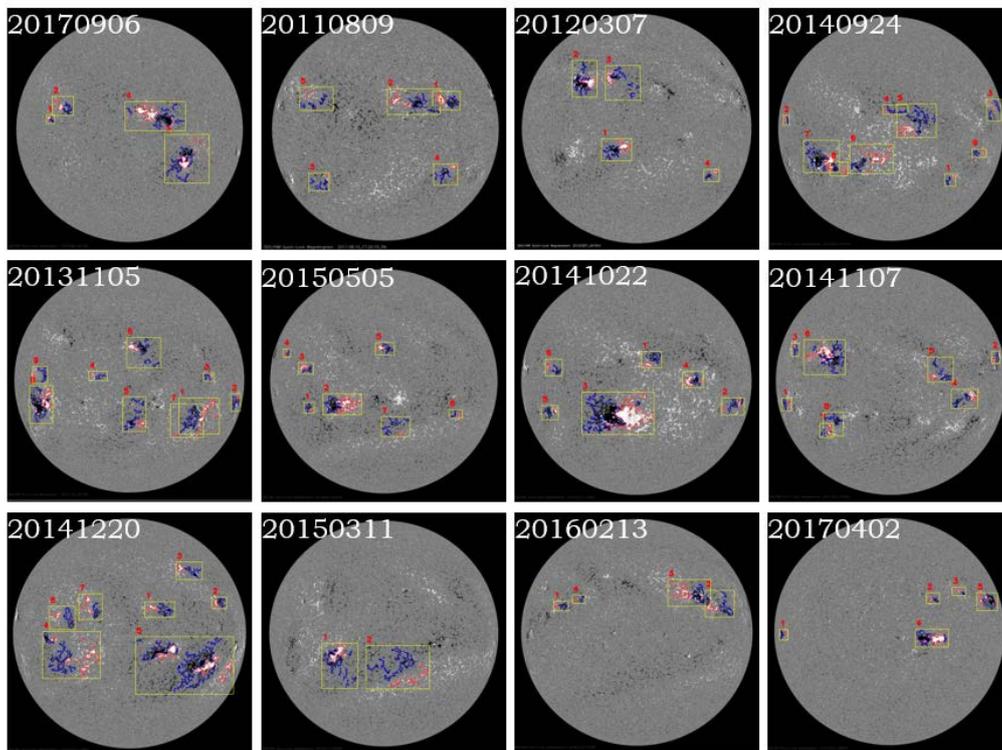
### 3.3. Pengujian Berdasarkan Matriks Evaluasi

Pengujian berdasarkan peristiwa *flare* menggunakan 12 citra HMI Magnetogram pada saat peristiwa *flare* kelas X dan M selama satelit SDO mengorbit atau setelah tahun 2010 dan peristiwa *flare* tersebut tidak terjadi di pinggiran piringan Matahari. Data historis peristiwa *flare* diperoleh dari <https://www.solarmonitor.org/>. Situs web tersebut menyediakan data historis daerah aktif berbasis penomoran NOAA dan kekuatan *flare* yang terjadi.

Tabel 1 merupakan hasil pengujian daerah aktif yang terdeteksi dan dievaluasi oleh matriks evaluasi *correctness*, *completeness*, dan *quality*. Persentase *correctness* menunjukkan ketepatan program dalam mendeteksi daerah aktif dibandingkan dengan jumlah yang terdeteksi, sedangkan *completeness* mengukur ketepatan program dalam mendeteksi daerah aktif berdasarkan data historis. Gambar 10 adalah hasil deteksi dan pengelompokan daerah aktif menggunakan algoritma ACC secara berturut-turut sesuai dengan urutan data pada Tabel 1.

Tabel 1. Hasil matriks evaluasi deteksi dan pengelompokan daerah aktif.

Tanggal (TTTTBBHH)	AC	ADC	ARC	Correctness	Completeness	Quality
20170906	4	4	4	100	100	100
20110809	3	5	3	60	100	60
20120307	4	4	4	100	100	100
20140924	7	9	7	78	100	78
20131105	7	9	7	78	100	78
20150505	5	7	5	72	100	72
20141022	4	6	5	67	80	57
20141107	4	8	6	50	80	45
20141220	5	7	8	71	63	50
20150311	1	2	2	50	50	33
20160213	4	4	4	100	100	100
20170402	4	5	4	80	100	80



Gambar 10. Hasil deteksi dan pengelompokan daerah aktif pada 12 citra HMI Magnetogram.

Program gagal mendeteksi 3 daerah aktif pada data 20141220. Selain itu, program hanya berhasil mendeteksi 1 dari 2 daerah aktif pada data 20150311. Hasil deteksi pada kedua data tersebut menyebabkan nilai perbandingan AC dan ADC yang kecil sehingga berpengaruh terhadap persentase *correctness*, *completeness*, dan *quality*, seperti yang terlihat pada Tabel 1.

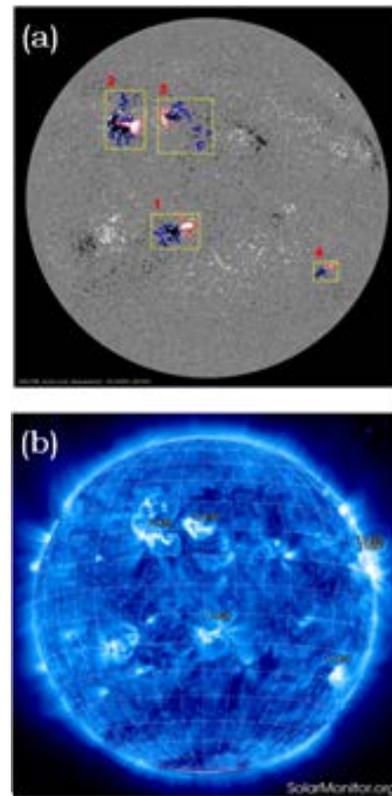
Pada data lainnya, program berhasil mendeteksi seluruh daerah aktif dengan tepat pada data 20170906, 20120307, dan 20160213 sehingga diperoleh persentase *correctness*, *completeness*, dan *quality* yang maksimal. Data 20140924, 20131105, 20150505, dan 20170402 memiliki tingkat *correctness* dan *quality* lebih dari 70%, dan tingkat *completeness* sebesar 100%.

Persentase *completeness* untuk data 20110809, 20141022, dan 20141107 di atas 80%, sedangkan untuk *correctness* dan *quality* masih di bawah 60%. Persentase *correctness* dan *quality* pada ketiga data tersebut dipengaruhi oleh perbandingan AC dan ADC yang timpang karena program mendeteksi daerah aktif lainnya yang tidak terdapat pada data historis atau yang dikenal dengan *over-specification*. *Over-specification* dari hasil program dapat dikatakan sebagai daerah dengan aktivitas magnetik rendah pada permukaan Matahari (*plage*) karena pasangan polaritas yang terdeteksi tidak tercatat sebagai daerah aktif pada penomoran NOAA.

### 3.4. Korelasi Peristiwa Flare dan Luas Daerah Aktif Pengujian Berdasarkan Matriks Evaluasi

Penelitian yang telah dilakukan oleh (Sulistiani, 2018) menjadikan luas daerah aktif sebagai masukan dalam memprediksi peristiwa *flare* menggunakan algoritma *random forest*. Penelitian tersebut menunjukkan bahwa luas daerah aktif memiliki korelasi dengan prakiraan daerah penghasil peristiwa *flare*. Hasil ekstraksi luas area terhadap 12 data pada penelitian ini akan dianalisis menggunakan korelasi sederhana (*bivariate correlation*) dengan kemunculan peristiwa *flare* pada daerah aktif. Gambar 11 merupakan contoh cara membandingkan hasil dengan data historis tanggal 20120307, dengan citra SWAP yang telah diberi keterangan penomoran daerah aktif NOAA

dan rekap peristiwa *flare* pada setiap daerah aktif.



Gambar 11. Perbandingan (a) citra tanggal 20120307 pada hasil program dan (b) citra yang telah diberikan penomoran daerah aktif NOAA dari solarmonitor.org.

Data solarmonitor.org mencatat bahwa peristiwa kelas *flare* C, M, dan X terjadi pada daerah aktif 11429 dan program berhasil mengidentifikasi daerah aktif tersebut sebagai daerah aktif nomor 2 pada Gambar 11(a). Daerah aktif tersebut memiliki luas daerah terbesar dibandingkan dengan daerah aktif lainnya, yaitu sebesar 3255 SMH.

Grafik pada Gambar 12 merupakan perbandingan luas daerah aktif pada setiap data. Penghasil peristiwa *flare* pada grafik ditunjukkan dengan keterangan tulisan kelas *flare*. Grafik menunjukkan luas daerah terbesar pada masing-masing data merupakan daerah penghasil *flare*, terkecuali untuk data pada tanggal 6 September 2017, dapat dilihat pada Gambar 12(a). Dengan demikian, dapat disimpulkan bahwa algoritma ACC berhasil mendeteksi daerah aktif dan ekstraksi parameter luas daerah aktif terbesar pada 12 citra HMI Magnetogram dan 11 di antaranya merupakan penghasil peristiwa *flare*.



Gambar 12. Grafik luas daerah aktif pada 12 citra yang membandingkan luas yang terdeteksi pada satu citra.

### 4. Kesimpulan

Penelitian ini telah berhasil mengembangkan aplikasi menggunakan teknik pengolahan citra digital untuk mendeteksi dan mengelompokkan daerah aktif yang terdiri dari area positif dan negatif pada citra HMI Magnetogram. Algoritma ACC mengelompokkan daerah aktif menggunakan data *centroid* area hitam dan putih yang terdeteksi secara dinamis berdasarkan jarak minimal dan perpotongan garis batas kontur area. Data *centroid* diperoleh dari proses *preprocessing* dimulai dari membaca masukan citra HMI Magnetogram dan mengonversikan ke dalam satuan piksel, konversi citra berwarna ke dalam citra abu-abu, menghilangkan derau pada citra menggunakan *median filter*, klasifikasi jenis area berdasarkan nilai batas, dan mencari informasi kontur pada setiap area. Kontur area hitam dan putih yang terdeteksi dikelompokkan ke dalam klaster berdasarkan jarak *centroid* dan perpotongan garis batas kontur area. Klaster area hitam dan putih yang berpasangan merupakan representasi dari suatu daerah aktif. Penentuan pasangan dilakukan berdasarkan perpotongan garis batas klaster area hitam dan putih.

Pengujian akurasi pengelompokan daerah aktif menggunakan matriks evaluasi *area-based accuracy* yang terdiri dari *correctness*, *completeness*, dan *quality* terhadap 12 data historis peristiwa *flare* kelas X dan M. Berdasarkan matriks evaluasi, penelitian ini telah berhasil mendeteksi daerah aktif dengan persentase *correctness*, *completeness*, dan *quality* yang berbeda pada setiap data uji. Tujuh data uji menunjukkan tingkat akurasi yang tinggi berdasarkan persentase *correctness*, *completeness* dan *quality*. Pada tiga data uji, algoritma ACC mendeteksi daerah aktif yang tidak tercatat sebagai daerah aktif berdasarkan penomoran NOAA atau yang biasa disebut *over-specification*.

Ekstraksi parameter luas daerah aktif pada penelitian ini menunjukkan bahwa luas daerah aktif terbesar pada satu citra merupakan daerah penghasil *flare*. Hasil ekstraksi luas daerah aktif pada penelitian ini diharapkan dapat digunakan oleh peneliti dengan bidang kepakaran terkait untuk memprediksi peristiwa *flare* ataupun menganalisis secara fisis.

## Ucapan Terima Kasih

Ucapan terima kasih disampaikan kepada tim SDO GSFC NASA <https://sdo.gsfc.nasa.gov/> yang telah menyediakan data HMI Magnetogram, dan tim <https://www.solarmonitor.org/> yang telah memberikan informasi *flare* secara terperinci.

## Pernyataan Kontribusi

P. Pertama dan P. Kedua merupakan kontributor utama yang mengakusisi, mengolah data, melakukan analisis, merancang penelitian, dan menulis sebagian besar makalah ini.

## Rujukan

- Ashamari, O., Qahwaji, R., Ipson, S., Schöll, M., Nibouche, O., & Haberleiter, M. (2015). Identification of photospheric activity features from SOHO/MDI data using the ASAP tool. *Journal of Space Weather and Space Climate*, 5, A15.
- Bobra, M.G., Sun, X., Hoeksema, J.T., Turmon, M., Liu, Y., Hayashi, K., & Leka, K.D. (2014). The Helioseismic and Magnetic Imager (HMI) vector magnetic field pipeline: SHARPs—space-weather HMI active region patches. *Solar Physics*, 289(9), 3549-3578.
- Cai, L., Shi, W., Miao, Z., & Hao, M. (2018). Accuracy assessment measures for object extraction from remote sensing images. *Remote Sensing*, 10(2), 303.
- Duffet-Smith, P. (1988). *Practical astronomy with your calculator*, 3<sup>rd</sup> ed. Cambridge, Univ.Press.
- Georgoulis, M.K. (2008). Magnetic complexity in eruptive solar active regions and associated eruption parameters. *Geophysical research letters*, 35(6).
- Guglielmino, S.L., Pillet, V.M., Bonet, J.A., del Toro Iniesta, J.C., Rubio, L.B., & *et al.* (2012). The frontier between small-scale bipoles and ephemeral regions in the solar photosphere: emergence and decay of an intermediate-scale bipole observed with SUNRISE/IMaX. *The Astrophysical Journal*, 745(2), 160.
- Hidayatullah, P. (2017). *Pengolahan Citra Digital Teori dan Aplikasi Nyata*. Bandung, Informatika.
- Higgins, P.A., Gallagher, P.T., McAteer, R.J., & Bloomfield, D.S. (2011). Solar magnetic feature detection and tracking for space weather monitoring. *Advances in Space Research*, 47(12), 2105-2117.
- Howse, J. (2013). *OpenCV computer vision with python*, Packt Publishing Ltd.
- Kaushik, M. & Mathur, B. (2014). Comparative study of K-means and hierarchical clustering techniques. *International journal of software and hardware research in engineering*, 2(6), 93-98.
- Mackay, D.H., Green, L.M., & Van Ballegooijen, A. (2011). Modeling the dispersal of an active region: quantifying energy input into the corona. *The Astrophysical Journal*, 729(2), 97.
- Meadows, P.J. (2020). Remeasurement of Solar Observing Optical Network sunspot areas. *Monthly Notices of the Royal Astronomical Society*, Volume 497, Issue 1, pp.1110-1114, doi: 10.1093/mnras/staa2007.
- Schrijver, C.J., Kauristie, K., Aylward, A.D., Denardini, C.M., Gibson, & *et al.* (2015). Understanding space weather to shield society: A global road map for 2015–2025 commissioned by COSPAR and ILWS. *Advances in Space Research*, 55(12), 2745-2807.
- Smart W.M. (1971). *Text-book on Spherical Astronomy*. Cambridge Univ. Press, Cambridge.
- Sreeja, V. (2016). Impact and mitigation of space weather effects on GNSS receiver performance. *Geoscience Letters*, 3(1), 24.
- Steward, G., Lobzin, V., Cairns, I.H., Li, B., & Neudegg, D. (2017). Automatic recognition of complex magnetic regions on the Sun in SDO magnetogram images and prediction of flares: Techniques and results for the revised flare prediction program Flarecast. *Space Weather*, 15(9), 1151-1164.
- Sulistiani, S. & Dani, T. (2019). Prakiraan Flare Sinar-X Matahari Berdasarkan Evolusi Daerah Aktif. *Jurnal Sains Dirgantara*, 16(1), 23-32.
- Suzuki, S. (1985). Topological structural analysis of digitized binary images by border following. *Computer vision, graphics, and image processing*, 30(1), 32-46.

