

Klasifikasi *Plasmodium vivax* dari Digitalisasi Mikroskopis Sediaan Darah Tebal Menggunakan Kombinasi Ekstraksi Ciri Statistik Orde Ke-dua dan Metode *K-Nearest Neighbor (K-NN) Classifier*

Niken Kartika Imania SP¹, T. Sutojo²

Abstract - Malaria is an infectious disease caused by parasites of the genus *Plasmodium*. There are four species causing malaria and each species cause different types of malaria which can be seen from the periodic interval of the onset of symptoms. Slow medical action on malaria patient can be fatal. Knowing the type and stage of malaria parasites is easier to do with the observation in the preparation of thin blood films than blood ones. In Indonesia, observation on the thick blood films is still widely practiced, whereas they more often used to discover the presence or absence of malaria parasites. Second order statistical feature is one of texture analysis methods. Feature extraction is closely related to texture analysis. The process of classification and image interpretation based on texture analysis is generally requires a feature extraction stage. *K-Nearest Neighbor (K-NN)* is one of classification techniques which algorithm is very simple but the performance is so reliable. Accuracy of the classification of the non-parasite image, *Plasmodium vivax* malaria parasite thropozoite, *Plasmodium vivax* schizont, and *Plasmodium vivax* gametocyte using combined of second order statistical feature extraction and *K-NN* reaches 95% when the value of *k* is 5 and 7. It is proved that the combined of this two methods can classify the type and stage of malaria parasites in thick blood films.

Keywords: *K-Nearest Neighbor (K-NN) Classifier, Malaria, Plasmodium vivax, Second Order Statistical Feature Extraction, Thick Blood Films.*

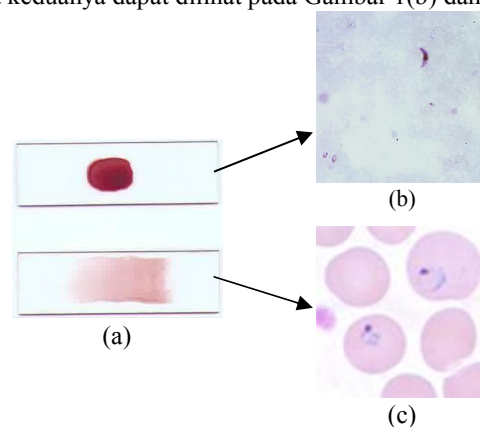
I. PENDAHULUAN

Malaria adalah salah satu penyakit yang masih menjadi perhatian besar di daerah tropis dan subtropis, termasuk salah satunya adalah Indonesia. Penyakit menular yang menginfeksi sel darah merah ini disebabkan oleh parasit protozoa dari genus *plasmodium* dan ditularkan melalui gigitan nyamuk *Anopheles* sp. betina. Di antara 170 jenis *plasmodium*, ada empat penyebab malaria pada manusia, yaitu *Plasmodium falciparum*, *Plasmodium malariae*, *Plasmodium ovale*, dan *Plasmodium vivax*. Di Indonesia, spesies *plasmodium* yang sering ditemui adalah *Plasmodium falciparum* (penyebab malaria tropika) dan *Plasmodium vivax* (penyebab malaria tertiana).

Data Kementerian Kesehatan menyebutkan bahwa Provinsi Papua, Papua Barat, dan Nusa Tenggara Timur masih menjadi provinsi dengan kasus malaria paling besar. Papua memiliki persentase kasus terbesar yakni 42,65 %, sementara Papua Barat sebesar 38,44 % dan NTT sebesar 16,37 % dari 93,2 % konfirmasi kasus malaria yang ada di Indonesia sepanjang tahun 2013 [2].

Salah satu teknik untuk diagnosa pasti malaria yang tidak

hanya dapat menemukan jenis tetapi juga stadium *plasmodium* adalah dengan membaca mikroskopis sediaan darah yang dapat dibuat dalam dua bentuk, yaitu sediaan darah tipis / sediaan apus darah dan sediaan darah tebal / sediaan tetes tebal [14], dimana perbedaan keduanya dapat dilihat pada Gambar 1(a) dan pengamatan yang dilakukan pada keduanya dapat dilihat pada Gambar 1(b) dan (c).



Gambar 1. (a) Sediaan darah tebal (atas) dan tipis (bawah) (b) Citra Sediaan Darah Tebal (c) Citra Sediaan Darah Tipis

Sediaan darah tipis lebih sering digunakan untuk menentukan jenis dan fase parasit [5] serta jenis obat yang akan diberikan kepada penderita malaria [7]. Eritrosit pada sediaan darah tipis masih utuh, sehingga bentuk dan morfologi parasit masih utuh dan sempurna [10][7]. Struktur parasit pada sediaan darah tipis yang terlihat lebih jelas inilah yang menyebabkan spesies *plasmodium* dan stadium (fase)nya lebih mudah untuk diidentifikasi [7].

Di Indonesia, pengamatan yang banyak digunakan adalah pengamatan pada sediaan darah tebal. Pengamatan mikroskopis pada sediaan darah tebal lebih sering digunakan untuk sebatas menemukan ada atau tidaknya parasit malaria [5].

Berdasarkan fakta tersebut, dibutuhkan metode untuk menentukan jenis dan fase parasit malaria pada sediaan darah tebal [5] karena diagnosis malaria yang tepat adalah langkah awal dalam memberikan pengobatan yang tepat dan cepat sehingga komplikasi dapat dihindari [7].

Penelitian dalam mengklasifikasi jenis dan fase parasit malaria telah dilakukan dengan berbagai metode, diantaranya menggunakan *Artificial Neural Network* (ANN), *Active Contour* dan *Support Vector Machine*, sertakombinasi antara *Principal Component Analysis* (PCA) dan ekstraksi ciri statistik orde ke-dua [5][6][15].

II. METODE YANG DIUSULKAN

Metode yang penulis gunakan untuk mengolah citra dalam penelitian tugas akhir ini adalah ekstraksi ciri statistik orde ke-dua, sedangkan untuk melakukan klasifikasi jenis dan fase parasit malaria, penulis menggunakan metode klasifikasi *K-Nearest Neighbor* (K-NN).

Ciri statistik orde ke-dua merupakan salah satu metode untuk analisis tekstur. Ekstraksi ciri erat kaitannya dengan analisis tekstur. Proses klasifikasi dan interpretasi citra dengan dasar analisis tekstur pada umumnya membutuhkan tahapan ekstraksi ciri [1]. *K-Nearest Neighbor* (K-NN) merupakan salah satu teknik klasifikasi yang algoritmanya sangat sederhana namun performanya sangat handal [4].

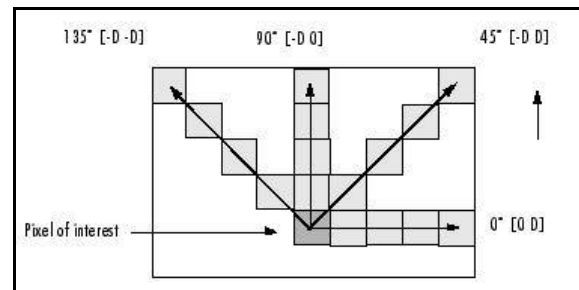
Berdasarkan tinjauan tersebut, penulis tertarik untuk melakukan penelitian terhadap mikroskopis sediaan darah tebal terinfeksi parasit malaria menggunakan ekstraksi ciri statistik orde ke-dua dan *K-Nearest Neighbor* untuk identifikasi *Plasmodium vivax*. Penelitian ini kemudian penulis beri judul “Identifikasi *Plasmodium vivax* dari Mikroskopis Sediaan Darah Tebal Menggunakan Ekstraksi Ciri Statistik Orde ke-Dua dan Metode *K-Nearest Neighbor* (K-NN) *Classifier*”.

A. Ekstraksi Ciri Statistik Orde Ke-dua

Ekstraksi ciri erat kaitannya dengan analisis tekstur. Sebuah citra berbasis analisis tekstur yang akan diklasifikasi umumnya membutuhkan tahapan ekstraksi ciri sebagai langkah awal untuk dilakukan kuantisasi karakteristik citra ke dalam sekelompok nilai ciri yang sesuai. Tinggi suatu orde menunjukkan tingkat kesukaran proses ekstraksinya.

Makin tinggi ordenya, makin tinggi pula tingkat kesukaran proses ekstraksinya. Ekstraksi ciri statistik orde ke-dua merupakan ekstraksi ciri pada citra dengan mempertimbangkan hubungan antara dua piksel sebuah citra pada jarak dan orientasi sudut tertentu.

Salah satu teknik untuk memperoleh ciri statistik orde dua adalah dengan menghitung probabilitas hubungan ketetanggaan antara dua piksel pada jarak dan orientasi sudut tertentu. Pendekatan ini bekerja dengan membentuk sebuah matriks ko-okurensi atau sering disebut dengan *Gray Level Co-occurrence Matrix* (GLCM) dari data citra, dilanjutkan dengan menentukan ciri sebagai fungsi dari matriks antara tersebut. Ko-okurensi dalam hal ini adalah jumlah kejadian satu level nilai suatu piksel bertetangga dengan satu level nilai piksel lain yang terjadi secara bersamaan pada jarak d dan orientasi sudut θ tertentu, dimana jarak dinyatakan dalam piksel dan orientasi dinyatakan dalam derajat. Orientasi yang digunakan dalam penelitian ini adalah empat arah sudut dengan interval 45° , yaitu 0° , 45° , 90° , dan 135° .



Gambar 2. GLCM yang menggambarkan korelasi antara sepasang piksel dengan jarak dan arah tertentu [14]

Terdapat beberapa ciri/fitur pada citra sebagai pembeda antara citra satu dengan lainnya, diantaranya adalah *contrast*, *correlation*, *energy*, dan *homogeneity*.

1. Contrast

$$CON = \sum_i \sum_j |(i - j)|^2 P_d(i, j) \quad (2-1)$$

2. Correlation

$$COR = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) p(i, j)}{\sigma_i \sigma_j} \quad (2-2)$$

dimana

μ_i adalah nilai rata-rata elemen baris pada matriks $Pd\theta(i, j)$, dinyatakan dengan rumus

$$\mu_i = \sum_i \sum_j i p(i, j) \quad (2-3)$$

μ_j adalah nilai rata-rata elemen kolom pada matriks $Pd\theta(i, j)$, dinyatakan dengan rumus

$$(2-4)$$

$$\mu_j = \sum_i \sum_j p(i, j)$$

σ_i adalah nilai standar deviasi elemen baris pada matriks $Pd\theta(i, j)$, dinyatakan dengan rumus

$$\sigma_i = \sqrt{\sum_{i,j} (i - \mu_i)^2 p(i, j)} \quad (2-5)$$

σ_j adalah nilai standar deviasi elemen kolom pada matriks $Pd\theta(i, j)$, dinyatakan dengan rumus

$$\sigma_j = \sqrt{\sum_{i,j} (j - \mu_j)^2 p(i, j)} \quad (2-6)$$

3. Energy

$$ENG = \sum_{i,j} P_d^2(i, j) \quad (2-7)$$

4. Homogeneity

$$HOM = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (2-8)$$

B. K-Nearest Neighbor (K-NN) Classifier

K-Nearest Neighbor (K-NN) salah satu teknik untuk melakukan klasifikasi. K-NN termasuk ke dalam kategori *supervised learning*, yaitu suatu metode yang membutuhkan fase pembelajaran. Prinsip kerja algoritma ini adalah menghitung jarak suatu objek baru terhadap data pembelajaran lalu mengelompokkannya ke dalam kelas mayoritas yang jaraknya paling dekat dengan objek. Itulah sebab mengapa K-NN dikategorikan sebagai salah satu algoritma *supervised learning*. Sebuah objek dikelompokkan ke dalam kelas a jika kelas a merupakan kelas dengan atribut paling banyak yang ditemui pada k buah tetangga terdekat objek tersebut. Algoritma K-Nearest Neighbor adalah sebagai berikut [16] :

1. Tentukan nilai parameter k , yaitu jumlah tetangga terdekat.
2. Hitung jarak antara citra acuan dengan semua citra uji menggunakan jarak *euclidean*.
3. Urutkan jarak tersebut dan tetapkan tetangga terdekat berdasarkan jarak k minimum.
4. Periksa *output* atau label pada masing-masing kelas tetangga terdekat.

5. Kelompokkan citra baru (citra uji) tersebut ke dalam kelas mayoritas dari tetangga terdekat.

Jauh atau dekatnya tetangga dihitung berdasarkan jarak. Terdapat beberapa algoritma untuk menghitung jarak. Pada penelitian ini, penulis menggunakan jarak *euclidean* yang dijabarkan sebagai berikut :

$$d(x, y) = \sqrt{\sum_{i=1}^n (a_i - A_i)^2 + (b_i - B_i)^2 + (c_i - C_i)^2 + (d_i - D_i)^2} \quad (2-9)$$

dengan :

$d(x, y)$: jarak citra latih dengan citra uji

a adalah fitur ke- i dari citra x , b adalah fitur ke- i dari citra y , dan n adalah jumlah fitur.

III. IMPLEMENTASI

3.1. Prapengolahan

Pada tahap prapengolahan citra, semua citra mikroskopis sediaan darah tebal baik pada data latih maupun data uji dinormalisasi dengan cara memotong langsung objek bukan parasit dan parasit malaria kemudian melakukan *resize* pada potongan citra tersebut secara manual menjadi dimensi 60 x 60 piksel. Hasil pada tahap ini adalah citra bukan parasit dan parasit malaria dimensi 60 x 60 piksel.

3.2. Pengolahan Citra

Tahap pengolahan citra pada penelitian ini terdiri dari 3 subtahap. Subtahap 1 adalah mengubah citra RGB menjadi citra keabuan. Citra hasil tahap 1 diubah menjadi citra *grayscale* dengan 256 derajat keabuan yang merupakan *default* derajat keabuan citra *grayscale*. Guna mengurangi beban komputasi, citra tersebut diubah menjadi citra *grayscale* 8 derajat keabuan. Hasil pada subtahap ini adalah citra *grayscale* 8 derajat keabuan. Hasil dari subtahap 1 digunakan untuk membuat matriks ko-okurensi. Pada penelitian ini, matriks ko-okurensi dibuat dengan menghitung hubungan spasial pada citra *grayscale* 8 derajat keabuan dengan perhitungan sudut 0°, 45°, 90°, 135° dan jarak sebesar 1 piksel. Hasil subtahap 2 adalah empat matriks ko-okurensi tersebut. Pada subtahap 3, empat matriks ko-okurensi masing-masing digunakan untuk ekstraksi ciri citra. Ciri citra yang akan diekstrak dalam penelitian ini berjumlah empat, yaitu *contrast*, *correlation*, *energy*, dan *homogeneity*. Nilai-nilai ciri citra inilah yang digunakan sebagai dasar untuk melakukan tahap 3, yaitu klasifikasi.

3.3. Klasifikasi Menggunakan Teknik K-NN dengan Jarak *Euclidean*

Klasifikasi dalam penelitian ini menggunakan teknik K-NN berbasis jarak *euclidean*, yaitu menghitung jarak minimum antara citra uji terhadap semua citra latih setelah menetapkan tetangga terdekat. Citra akan teridentifikasi berdasarkan kelas mayoritas, yaitu kelas dengan anggota terbanyak dari tetangga terdekat.

3.4. Validasi dilakukan dengan menghitung nilai akurasi, yaitu membagi jumlah citra uji yang teridentifikasi benar dengan total citra uji kemudian mengalikannya dengan 100 % untuk mendapatkan persentase.

IV. HASIL DAN PEMBAHASAN

Tabel 1. Kelompok dan nama kelas

| Kelompok / Kelas | Nama Kelompok / Kelas |
|------------------|------------------------------------|
| 1 | Bukan Parasit |
| 2 | <i>Plasmodium vivax thropozoit</i> |
| 3 | <i>Plasmodium vivax schizont</i> |
| 4 | <i>Plasmodium vivax gametocyte</i> |

Klasifikasi terdiri dari empat kelas seperti yang terdapat pada tabel di atas. Pada penelitian ini, nilai k yang dipakai adalah 1,3,5,7, dan 9. Hasil pengujian dari proses klasifikasi yang telah dilakukan menggunakan K-NN dengan data latih dan data uji masing-masing 80 dan 40 untuk klasifikasi terhadap 4 (empat) kelas didapatkan nilai akurasi dari setiap nilai k seperti pada Tabel 2.

Tabel 2. Tingkat pengenalan sistem

| Nilai k | Jumlah Data Benar | Jumlah Data Salah | Tingkat Pengenalan (Akurasi) |
|-----------|-------------------|-------------------|------------------------------|
| 1 | 35 | 5 | 87,5 % |
| 3 | 37 | 3 | 92,5 % |
| 5 | 38 | 2 | 95 % |
| 7 | 38 | 2 | 95 % |
| 9 | 37 | 3 | 92,5 % |

Tabel 2 menunjukkan tingkat akurasi tertinggi diperoleh ketika nilai k adalah 5 dan 7, hal ini membuktikan bahwa gabungan kedua metode ini dapat mengklasifikasi jenis dan fase parasit malaria pada sediaan darah tebal.

V. PENUTUP

3.1. Kesimpulan :

Dari hasil penelitian yang telah dilakukan, maka dapat diambil kesimpulan sebagai berikut :

1. Empat ciri atau fitur yang diambil dari ROI citra

mikroskopis sediaan darah tebal dapat membedakan antara citra satu dengan lainnya.

2. Nilai akurasi yang diperoleh dari ekstraksi ciri statistik orde ke-dua dan teknik klasifikasi K-*Nearest Neighbor* (K-NN) *Classifier* pada penelitian ini menunjukkan bahwa kedua metode ini dapat mengidentifikasi serta mengelompokkan jenis dan fase parasit malaria pada sediaan darah tebal.

3.2. Saran

Berdasarkan tingkat pengenalan yang ditunjukkan penelitian ini, saran yang dapat penulis berikan adalah sebagai berikut :

1. Perlu dilakukan pemilihan data acuan secara optimal yang mampu mewakili representasi sebuah populasi, sehingga didapat hasil pengenalan yang lebih baik lagi, mengingat terdapat satu citra dimana dengan parameter (nilai k) yang berbeda-beda teridentifikasi sebagai citra yang berbeda dari kelas semestinya.
2. Perlu dikembangkan suatu metode yang mampu meminimalisir kesalahan pengenalan.

DAFTAR PUSTAKA

- [1] PRAKTIKUM PENGOLAHAN CITRA BIOMEDIKA. "Modul 3 – Analisis Tekstur". Imaging & Image Processing Research Group ITB.
- [2] Wilda, Z.F. "Inilah Alasan Angka Kasus Malaria di Indonesia Timur Masih Tinggi," 28 Februari 2014. [Online]. Available: detik.com/health/read/2014/02/28/180434/2511903/763/inilah-alasan-angka-kasus-malaria-di-indonesia-timur-masih-tinggi. [Accessed 19 Oktober 2014].
- [3] Berita Negara Republik Indonesia Tahun 2012. "Pedoman Penatalaksanaan Kasus Malaria". Jakarta, 2012. [Online]. Available: http://pppl.depkes.go.id/_asset/_download/Pedoman%20Penatalaksanaan%20kasus%20Malaria%202012.pdf [Accessed 11 November 2014].
- [4] Nugroho, A.S. "k-Nearest Neighbor Classifier," 26 Januari 2007. [Online]. Available: <https://asnugroho.wordpress.com/2007/01/26/k-nearest-neighbor-classifier/> [Accessed 19 November 2014].
- [5] Farah Z.R., I Ketut E.P., Mauridhi H.P., "Identifikasi Parasit Malaria pada Thick-Blood Film Menggunakan Artificial Neural Network", Jurusan Teknik Elektro, Institut Teknologi Sepuluh November, Surabaya, 2013.
- [6] Andi Kusuma Indrawan, "Identifikasi Fase Plasmodium Falciparum Menggunakan Active Contour dan Support Vector Machine", Jurusan Teknik Elektro, Fakultas Teknologi Industri, Institut Teknologi Sepuluh November, Surabaya, 2011.
- [7] Nugraha, A.C. "Laporan Magang Diagnosis Malaria dengan Pemeriksaan Mikroskopis di Balai Litbang P2B2 Banjarnegara", 2012.
- [8] CDC Centers for Disease Control and Prevention. "Biology," 09 November 2012. [Online].

- Available:<http://www.cdc.gov/malaria/about/biology/>. [Accessed 10 Desember 2014].
- [9] Kasatnakes. "Tentang Plasmodium Vivax," 21 Desember 2012. [Online]. Available: <http://www.katsanakes.com/2012/12/plasmodium-vivax.html>. [Accessed 10 Desember 2014].
- [10] dr. Ayda Rahmad, MS, Sp.ParK and Drs. Purnomo, SKM, Atlas Diagnostik Malaria, Jakarta : Penerbit EGC, 2011.
- [11] CDC Centers for Disease Control and Prevention. "Glossary," 09 November 2012. [Online]. Available: <http://www.cdc.gov/malaria/glossary.html>. [Accessed 10 Desember 2014].
- [12] Prasetyo, E. Pengolahan Citra Digital dan Aplikasinya Menggunakan Matlab, Yogyakarta : Penerbit Andi, 2011.
- [13] Sumanto, D. "Pembuatan Sediaan Malaria," 25 Februari 2012. [Online]. Available: didik.dosen.unimus.ac.id/2012/02/25/pembuatan-sediaan-malaria/ [Accessed 17 Desember 2014].
- [14] Darmawan, M. "Image Processing," 25 Februari 2011. [Online]. Available: <https://martendarmawan.wordpress.com/category/image-processing/>. [Accessed 17 Desember 2014].
- [15] Iis Hamsir A.W., Adhi Susanto, P. Insap S., Maesadji T., "Principal Component Analysis Combined with Second Order Statistical Feature Method for Malaria Parasite Classification", Fakultas Teknik Elektro dan Teknologi Informasi, Unviversitas Gadjah Mada, Yogyakarta, 2014.
- [16] Teknomo, K. "K-Nearest Neighbors Tutorial," 2015. [Online]. Available: <http://people.revoledu.com/kardi/tutorial/KNN/>. [Accessed 12 Januari 2015].
- [17] R. Munir, "www.stei.itb.ac.id," [Online]. Available: <http://informatika.stei.itb.ac.id/~rinaldi.munir/Buku/Pengolahan%20Citra%20Digital/Bab-1Pengantar%20Pengolahan%20Citra.pdf>. [Accessed 5 Oktober 2014].
- [18] R. Munir, "www.stei.itb.ac.id," [Online]. Available: <http://informatika.stei.itb.ac.id/~rinaldi.munir/Buku/Pengolahan%20Citra%20Digital/Bab-2Pembentukan%20Citra.pdf>. [Accessed 5 Oktober 2014].
- [19] R. Munir, "www.stei.itb.ac.id," [Online]. Available: <http://informatika.stei.itb.ac.id/~rinaldi.munir/Buku/Pengolahan%20Citra%20Digital/Bab-11Citra%20Biner.pdf>. [Accessed 5 Oktober 2014].
- [20] T. Sutojo, Teori dan Aplikasi Aljabar Linier & Matrik, Yogyakarta: Penerbit ANDI, 2010.
- [21] Yudhistira Ganis K., Imam Santoso, and R. Rizal Isnanto, "Klasifikasi Citra Dengan Matriks Ko-Okurensi Aras Keabuan (Gray Level Co-occurrence Matrix-GLCM) Pada Lima Kelas Biji-Bijian", Jurusan Teknik Elektro, Universitas Diponegoro, Semarang.
- [22] SoftScientists. "Algoritma GLCM – Analisis Tekstur Citra Digital," 01 Januari 2015. [Online]. Available: www.softscience.web.id/2014/04/pengolahan-citra-digital-dengan-c.html?m=1. [Accessed 10 Januari 2015].
- [23] Wikipedia. "Akurasi dan Presisi," 05 April 2013. [Online]. Available: id.m.wikipedia.org/wiki/Akurasi_dan_presisi. [Accessed 17 Desember 2014].