



K-Means *Clustering* Berdasarkan Otsu *Thresholding* untuk Segmentasi Inti Leukosit

Chyntia Raras Ajeng Widiawati¹, Wiga Maulana Baihaqi², Tegar Insani³

^{1,2,3} Program Studi Teknologi Informasi, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

¹chyntiaraw@amikompurwokerto.ac.id, ²wiga@amikompurwokerto.ac.id*, ³inzani72@gmail.com

Abstract

White blood cells function as the human immune system, and help defend the body against viruses. In clinical practice, identification and counting of white blood cells in blood smears is often used to diagnose many diseases such as infection, inflammation, malignancy, leukemia. In the past, examination of blood smears was very complex, manual tasks were tedious and time-consuming. This research proposes the k-means clustering algorithm to separate white blood cells from other parts. However, k-means clustering has a weakness that is when determining the initial prototype values, so the otsu thresholding method is used to determine the threshold by utilizing global values, then proceed with morphological operations to refine the segmentation image. The results of segmentation are measured by the Positive Predictive Value (PPV) and Negative Positive Value (NPV) parameters. The results obtained prove that the use of otsu thresholding and morphological operations significantly increase the value of PPV compared to the value of PPV that does not use otsu thresholding. Whereas the NPV value increased but not significantly.

Keywords: segmentation, white blood cells, k-means clustering, otsu thresholding, morphological operations

Abstrak

Leukosit berfungsi sebagai sistem kekebalan tubuh manusia, dan membantu mempertahankan tubuh melawan virus. Dalam praktek klinis, identifikasi dan penghitungan leukosit dalam apusan darah sering digunakan untuk mendiagnosis banyak penyakit seperti infeksi, peradangan, keganasan, leukemia. Di masa lalu, pemeriksaan apusan darah sangat kompleks, tugas manual yang membosankan dan menghabiskan waktu. Pada penelitian ini diusulkan algoritma pengelompokan k-means untuk memisahkan leukosit dengan bagian yang lainnya. Akan tetapi k-means clustering memiliki kelemahan yaitu ketika menentukan nilai-nilai prototipe awal, sehingga metode otsu thresholding digunakan untuk penentuan ambang batas dengan memanfaatkan nilai global, kemudian dilanjutkan dengan operasi morfologi untuk memperhalus citra segmentasi. Hasil segmentasi diukur dengan parameter Positive Predictive Value (PPV) dan Negative Positive Value (NPV). Hasil yang didapatkan terbukti bahwa penggunaan otsu thresholding dan operasi morfologi meningkatkan nilai PPV yang diperoleh secara signifikan dibandingkan dengan nilai PPV yang tidak menggunakan otsu thresholding. Sedangkan pada nilai NPV meningkat tetapi tidak secara signifikan.

Kata kunci: segmentasi, leukosit, k-means clustering, otsu thresholding, operasi morfologi

1. Pendahuluan

Sel darah dibuat di sumsum tulang dari sel induk. Darah melewati sumsum tulang dan mengambil sel-sel darah yang telah berkembang penuh untuk sirkulasi dalam darah. Sumsum tulang menghasilkan sebagian besar dari

semua sel darah, seperti trombosit, eritrosit, dan leukosit. Yang terakhir ini secara aktif berpartisipasi dalam sistem kekebalan tubuh manusia, dan membantu mempertahankan tubuh melawan virus [1].

Dalam praktek klinis, identifikasi dan penghitungan leukosit dalam apusan darah sering digunakan untuk mendiagnosis banyak penyakit seperti infeksi, peradangan, keganasan, leukemia. Di masa lalu, pemeriksaan apusan darah sangat kompleks, tugas manual yang membosankan dan menghabiskan waktu. Saat ini, dengan perkembangan yang cepat dari metode yang dibantu komputer, sistem analisis sel otomatis dapat mendukung analisis gambar yang lebih cepat dan lebih dapat direproduksi daripada analisis manual [2].

Sistem analisis sel otomatis umumnya meliputi empat langkah: akuisisi gambar, segmentasi sel, ekstraksi fitur, dan klasifikasi [3]–[5]. Segmentasi sel sering dianggap sebagai langkah yang paling penting dan penting dalam proses, karena secara langsung mempengaruhi keakuratan dan kompleksitas waktu dari langkah-langkah selanjutnya.

Namun, segmentasi leukosit yang cepat dan kuat masih menjadi sebuah tantangan bagi peneliti. Alasannya ada tiga. Pertama, gambar hapusan darah asli berbeda secara signifikan karena teknik pewarnaan yang berbeda dan kondisi pencahayaan. Kedua, variasi mungkin ada bahkan dalam citra smear yang sama karena berbagai jenis leukosit. Misalnya, leukosit biasanya dapat diklasifikasikan menjadi lima jenis (yaitu, limfosit, monosit, neutrofil, eosinofil, dan basofil) dan berbagai jenis eritrosit yang diwarnai dengan teknik yang sama menampilkan berbagai warna [6]. Ketiga, leukosit sering menyatu dengan eritrosit, menyebabkan bentuk leukosit yang tidak teratur, dan batas-batas antara sel-sel yang menyentuh menjadi kabur.

Berbagai metode segmentasi sel otomatis telah dikembangkan. Kebanyakan dari mereka adalah metode berbasis pembelajaran, yang umumnya dapat diklasifikasikan sebagai metode yang diawasi dan tidak diawasi. Metode yang diawasi, merumuskan masalah segmentasi sebagai klasifikasi multi-kelas. Misalnya, setiap piksel dapat diklasifikasikan ke dalam CROI dan non-CROI. Klasifikasi seperti, K-nn [7], Convolutional Neural Network [8], Support Vector Machine (SVM) [9], [10], telah digunakan, dimana pengklasifikasi biasanya dilatih pada gambar pelatihan yang diberi label secara manual. Efektivitasnya sangat tergantung pada kondisi pencitraan dan apakah fitur yang diekstrak dapat membedakan CROI dari non-CROI [11]. Dengan demikian, beberapa peneliti mencoba untuk mengekstraksi fitur warna yang lebih efektif untuk mengurangi masalah ini, misalnya, menggunakan ruang warna [12] atau ruang warna diskriminatif (MDC) [13] mengadopsi deskriptor skala-invarian fitur transformasi (SIFT) [11]. Selain menggunakan pengklasifikasi tradisional, jaringan convolutional yang mendalam [14]

juga baru-baru ini diterapkan dalam segmentasi gambar sel.

Metode pembelajaran tanpa pengawasan, seperti, K-means [15], fuzzy c-means (FCM) [16], di mana data pelatihan berlabel manual tidak diperlukan, memiliki kemampuan beradaptasi yang lebih kuat untuk berbagai kondisi daripada metode pembelajaran yang diawasi. Mereka efektif ketika warna seluruh sel berbeda dari yang lain. Namun, mereka tidak berkinerja baik ketika CROI memiliki variasi warna yang signifikan. Karena WBC memiliki warna yang sama dengan eritrosit dan sering melekat pada eritrosit, banyak parameter perlu disesuaikan kasus per kasus untuk menghindari segmentasi yang berlebihan dan kurang. Untuk mengatasi masalah ini, beberapa metode segmentasi sel berbasis clustering pertama-tama memisahkan latar belakang dari latar depan (termasuk CROI dan eritrosit) dan kemudian menghapus eritrosit dengan metode berbasis bentuk, menurut fakta bahwa sel-sel biasanya memiliki bentuk cembung. Ada terutama tiga jenis metode berbasis bentuk: Daerah aliran sungai berbasis [17], operasi dasar morfologi [18] dan analisis concavity berbasis kontur [19]. Metode analisis concavity, yang kurang sensitif terhadap ukuran sel dibandingkan dua metode sebelumnya, telah menarik lebih banyak perhatian. Namun, informasi intensitas abu-abu tidak dipertimbangkan dalam metode berbasis bentuk dan mereka mungkin gagal ketika bentuk sel sangat tidak teratur.

Dalam penelitian ini, kami mengusulkan pendekatan pembelajaran *unsupervised* untuk mencapai segmentasi inti leukosit yang cepat dan kuat serta berfokus pada segmentasi area inti sel. beberapa penelitian telah menggunakan algoritma k-means clustering untuk proses segmentasi leukosit atau inti leukosit, algoritma k-means dapat memisahkan dengan baik antara leukosit dan eritrosit atau memisahkan inti leukosit dengan bagian yang lainnya pada leukosit [15], [20]–[22]. Penelitian [20] menggunakan kombinasi peningkatan lokal menggunakan statistik histogram, pengelompokan k-means yang dimodifikasi, normalisasi, konversi ke citra biner menggunakan ambang batas global yang sesuai, penghilangan pulau dan pengisian lubang berdasarkan teknik pengisian benih, dan algoritma lokalisasi inti untuk segmentasi sel darah putih. Pada penelitian [21], segmentasi dilakukan menggunakan analisis cluster dengan metode K-Means pada citra AML yang memiliki variasi staining beragam. Sedangkan penelitian [22] telah berhasil membandingkan pengaruh dari tiga metode peningkatan kontras yaitu transformasi top hat dan bottom hat, linear contrast stretching dan histogram citra berbasis logika fuzzy sebagai tahap praproses untuk segmentasi K-means clustering sitoplasma sel darah putih menggunakan 15 citra. sampel darah model warna RGB, HSV dan Lab.

Masalah utama dari algoritma ini adalah ketergantungannya pada nilai-nilai prototipe awal. Jika prototipe awal tidak dipilih dengan hati-hati, perhitungan akan berjalan pada kesempatan konvergensi ke minimum lokal, bukan minimum global. Dengan demikian, pemilihan prototipe inisialisasi yang tepat memiliki dampak signifikan pada kinerja K-Means. Sehingga perlu adanya teknik penentuan ambang batas dengan memanfaatkan nilai global, salah satu metode yang dapat digunakan adalah otsu *thresholding*.

Sehingga pada penelitian ini metode yang diusulkan adalah k-means clustering untuk proses pemisahan inti sel dengan bagian yang lain, sedangkan otsu *thresholding* digunakan untuk binerisasi citra berdasarkan ambang batas global, dilanjutkan dengan operasi morfologi untuk memperhalus citra hasil segmentasi inti sel.

2. Metode Penelitian

Metodologi pada penelitian ini secara umum terdiri dari tiga tahapan. Tahapan umum tersebut diantaranya adalah dataset, pra-pengolahan, segmentasi citra, dan evaluasi segmentasi. Pada tahap pra-pengolahan sampai proses segmentasi menggunakan perangkat lunak Matlab R2015b 32 bit, sedangkan pada tahap evaluasi uji signifikansi perbedaan performa nilai PPV dan NPV menggunakan perangkat lunak Microsoft Excel.

2.1. Dataset

Dataset leukosit yang diperoleh dari Jiangxi Tecom Science Corporation, Cina. Dataset tersebut terdiri dari 300 citra dengan ukuran 120×120 dan kedalaman warnanya 24 bit. Citra diambil dengan kamera mikroskop optik Motic Moticam Pro 252A dengan mikroskop fokus otomatis N800-D, dan apusan darah diproses dengan *reagen* hematologi yang baru dikembangkan untuk pewarnaan eritrosit yang cepat. Latar belakang keseluruhan dari sebagian besar citra dataset terlihat kuning.

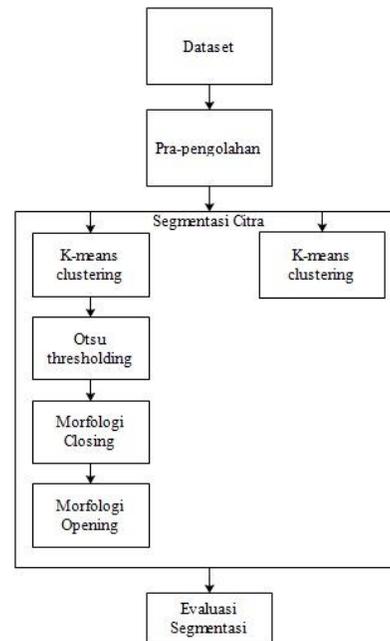
2.2. Pra-pengolahan

Pada tahap pra-pengolahan citra hasil akuisisi mengalami konversi citra dari RGB ke L^*a^*b . Hal tersebut berguna untuk memudahkan langkah berikutnya yaitu tahap *clustering* yang merupakan salah satu metode dalam tahap segmentasi.

2.3. Segmentasi Citra

Pada tahap ini terdiri beberapa metode yang digunakan. Tahapan tersebut diantaranya adalah *k-means clustering* yang bertujuan untuk mengelompokkan objek pada citra masukkan kedalam beberapa klaster berdasarkan nilai kedekatannya. Selanjutnya metode *Otsu thresholding* digunakan untuk menentukan ambang batas dari citra hasil

clustering, sehingga citra hasil *clustering* akan diubah menjadi citra biner. Hal tersebut bertujuan untuk memudahkan pemisahan inti dari leukosit dengan artefak yang lain. Kemudian citra biner tersebut akan mengalami dua kali operasi morfologi yaitu morfologi *closing* dan morfologi *opening*. Morfologi *closing* bertujuan untuk menyempurnakan hasil citra biner dengan cara mengisi bagian objek yang kosong. Kemudian dilanjutkan dengan operasi morfologi *opening*.



Gambar 1. Alur Metode Usulan

K-Mean adalah algoritma pembelajaran tanpa pengawasan untuk menyelesaikan masalah *clustering* [15], [22]. Lebih cepat dibandingkan dengan algoritma C-mean fuzzy. Algoritma ini mengambil jarak sebagai penghitungan parameter ketidaksamaan. Ini membagi objek p menjadi cluster k di mana setiap objek adalah bagian dari cluster dengan rata-rata terdekat. Dalam deteksi eritrosit, pengelompokan citra apusan darah dilakukan berdasarkan informasi warna eritrosit, leukosit dan plasma. Metode ini benar-benar memberikan kelompok yang berbeda dari perbedaan yang kuat. Tujuan utama dari algoritma pengelompokan ini adalah untuk meminimalkan jarak objek kuadrat dari pusat cluster, dan objek yang ditugaskan ke cluster yang memiliki jarak kuadrat Euclidean minimum. Dimana d_{ij} merupakan jarak antara objek i dengan j , x_{ik} merupakan nilai objek i pada variable ke- k , x_{jk} merupakan nilai objek j pada variable ke- k , dan p merupakan banyaknya variabel yang diamati.

$$d_{ij} = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (1)$$

Otsu merupakan teknik segmentasi paling sederhana dan paling efisien yang memisahkan objek dari latar belakang, mentransmisikan citra abu-abu atau warna menjadi citra biner di mana semua piksel dipartisi menjadi dua wilayah terpisah, satu berisi latar depan dan latar belakang lainnya [23]. Berbagai teknik *thresholding* telah digunakan oleh banyak peneliti seperti global, lokal dan *multiple thresholding* sesuai dengan kebutuhan mereka. Dalam global *thresholding* seluruh gambar di-*threshold* dengan nilai ambang tunggal T sedangkan partisi *thresholding* lokal seluruh gambar asli menjadi sub gambar dan setiap sub gambar memiliki nilai ambang tertentu.

Global thresholding memberikan kinerja yang baik jika ada pemisahan halus antara leukosit dan komponen darah lainnya di latar belakang.

$$I_{bin}(x, y) = \begin{cases} 1, & \text{if } I_{gray}(x, y) \geq T \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

Dalam percobaan kami, teknik ambang global *Otsu* telah digunakan dan algoritme ini cepat karena langsung beroperasi pada histogram tingkat abu-abu; di mana piksel gambar dipartisi ke dalam dua kelas *foreground* dan *background* yaitu, C_0 dan C_1 pada tingkat abu-abu t yang meminimalkan jumlah bobot varian kelas *foreground* dan piksel latar belakang dalam kelas untuk menetapkan ambang batas optimum t , yang merupakan sama dengan memaksimalkan varian antar kelas. Varians dalam kelas didefinisikan sebagai jumlah tertimbang dari kedua kluster seperti yang ditunjukkan pada Persamaan 3.

$$\sigma_p^2(t) = p_1(t)\sigma_1^2 + p_2(t)\sigma_2^2 \quad (3)$$

p_i adalah probabilitas dan σ_i^2 adalah varian dari kedua kelas di mana

$$p_1(t) = \sum_{i=0}^{T-1} q_i \quad (4)$$

$$p_2(t) = \sum_{i=T}^{N-1} q_i \quad (5)$$

dan rentang level intensitas diberikan oleh $[0, N-1]$.

2.3 Morfologi Opening dan Closing

Pada bagian ini, kami fokus pada operasi morfologis yang digunakan dalam pemrosesan gambar. Erosi, pelebaran, pembukaan, dan penutupan adalah empat operasi morfologis yang telah kami gunakan dalam penelitian kami [24]. Kami telah menggunakan operasi morfologis untuk menghilangkan benda yang tidak diinginkan seperti trombosit dan komponen darah lainnya dari gambar. Secara umum, *morphing matematika* digunakan di banyak bidang seperti segmentasi, peningkatan, analisis tekstur,

bentuk, ukuran, cembung *cull*, dan banyak lagi. Dalam penelitian kami, operasi morfologis digunakan sebagai langkah terakhir untuk menghaluskan inti sel. Misalkan kita memiliki gambar P dan elemen penataan Q , maka operasi dilambangkan pada Persamaan 6 dan 7.

$$\text{Erosin } P \ominus Q = \bigcap_{e \in Q} P_{-e} = \{e | (Q)_e \subseteq P\} \quad (6)$$

$$\text{Dilation } P \oplus Q = \bigcup_{e \in Q} P_e = \{e | (\hat{Q})_e \cap P \neq \emptyset\} \quad (7)$$

Dua operasi *morphing* lainnya membuka dan menutup yang berasal dari operasi erosi dan pelebaran. Bukaan berasal dari erosi gambar P oleh elemen penataan Q yang diikuti oleh pelebaran. Bukaan menghaluskan kontur objek, menghilangkan *isthmus* yang sempit, dan tonjolan tipis. Penutupan berasal dari pelebaran yang diikuti oleh erosi yang digunakan untuk menghaluskan kontur tetapi juga menghilangkan atau mengisi lubang-lubang kecil, menggabungkan celah-celah sempit dan jurang panjang tipis dalam kontur.

$$\text{Opening } P \circ Q = (P \ominus Q) \oplus Q \quad (8)$$

$$\text{Closing } P \bullet Q = (P \oplus Q) \ominus Q \quad (9)$$

2.4. Evaluasi Segmentasi

Citra hasil segmentasi kemudian akan dievaluasi dengan menggunakan citra *ground-truth*. Hasil evaluasi terhadap piksel citra yang bersesuaian dengan piksel pada citra *ground-truth* akan dievaluasi dengan menggunakan parameter pengukuran (PPV) dan (NPV). Nilai PPV digunakan karena PPV menunjukkan seberapa banyak persentase area yang benar dan berhasil diperoleh algoritma segmentasi dari area yang seharusnya. Sedangkan nilai NPV digunakan karena NPV menunjukkan seberapa banyak persentase area yang salah (bukan inti leukosit) dan berhasil diidentifikasi oleh algoritma segmentasi sebagai latar belakang maupun artefak lain. Semakin tinggi nilai PPV dan NPV, maka algoritma segmentasi yang digunakan semakin mendekati hasil segmentasi yang seharusnya.

Tingkat kebenaran suatu metode dapat dievaluasi dengan cara menghitung *true positive* (TP), *true negative* (TN), *false positive* (FP) dan *false negative* (FN). Keempat jenis data ini membentuk *confusion matrix* yang ditunjukkan pada. Dari keempat jenis elemen *confusion matrix* dapat ditarik beberapa parameter evaluasi, seperti akurasi, presisi, *recall*, *Fscore*, spesifisitas dan AUC [25], [26]. Untuk mengetahui tingkat keberhasilan metode segmentasi yang diusulkan, maka perlu dilakukan evaluasi terhadap hasil segmentasi yang diperoleh. Evaluasi terhadap hasil segmentasi diukur melalui beberapa

parameter pengukuran diantaranya *positive predictive value* (PPV) dan *negative predictive value* (NPV). Adapun cara perhitungan kedua parameter pengukuran tersebut dijelaskan pada persamaan (10) dan (11) berikut ini.

$$PPV = \frac{TP}{TP + FP} \quad (10)$$

$$NPV = \frac{TN}{TN + FN} \quad (11)$$

Untuk mengetahui tingkat signifikansi perbedaan antara performa segmentasi yang dihasilkan oleh metode *k-means clustering* dengan kombinasi antara *k-means clustering* dan *otsu thresholding*, perlu dilakukan uji *paired t-test*. Di bawah ini langkah-langkah dalam uji t-test :

1. Hitung $m_d = mx - my$, (m_d adalah perbedaan kedua *mean*)
2. Hitung $\sigma_d^2 = \frac{\sigma_x^2}{k} - \frac{\sigma_y^2}{l}$, (σ_d^2 adalah perbedaan kedua *variance*)

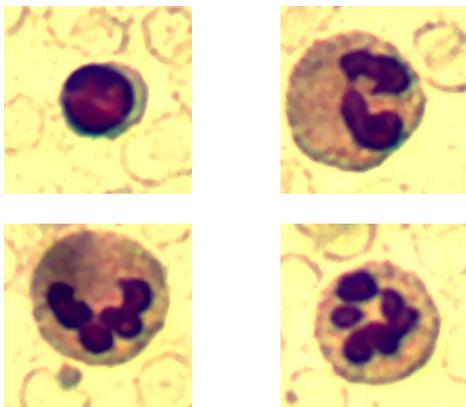
$$t = \frac{m_d}{\sqrt{\sigma_d^2 / k}}$$

3. Hitung $\sqrt{\sigma_d^2 / k}$, (Bentuk standar m_d disebut nilai t stat)
4. Nilai *Confidence limit z* dapat dibaca di tabel distribusi *student*.
5. Karena uji ini merupakan *two-tailed test*, maka tingkat s kemudian cari nilai z yang bersesuaian dengan tabel distribusi *student*.

3. Hasil dan Pembahasan

3.1. Dataset

Citra hasil akuisisi yang diperoleh dari Jiangxi Tecom Science Corporation.



Gambar 3. Citra Hasil Akuisisi

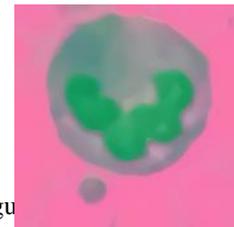
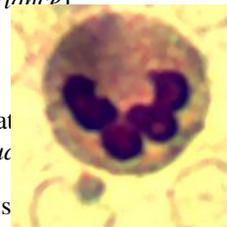
Berdasarkan Gambar 3 dapat dilihat bahwa citra hasil akuisisi terhadap leukosit memiliki inti sel yang berbeda-beda sesuai dengan jenis dari leukosit itu sendiri. Sehingga perlu dilakukan pelatihan terhadap algoritma yang akan

digunakan untuk mengenali ciri dari masing-masing jenis leukosit dengan mengambil informasi yang terkandung dari bentuk inti sel.

Citra hasil akuisisi memiliki format data .bmp dengan ruang warna RGB. Citra tersebut nantinya akan dikonversi ke ruang warna L*a*b pada tahap pra-pengolahan untuk mengekstrak inti sel dengan lebih baik. Hal tersebut dilakukan karena citra hasil akuisisi memiliki kontras yang cukup baik sehingga informasi warna menjadi salah satu fitur yang cukup menonjol untuk masing-masing sel. Pemaksimalan fitur warna dari citra hasil akuisisi akan membantu proses segmentasi dari bentuk inti sel pada tahap berikutnya.

3.2. Pra-pengolahan

Pada tahap pra-pengolahan, citra hasil akuisisi yang semula RGB dikonversi menjadi citra L*a*b untuk memudahkan tahap clustering.

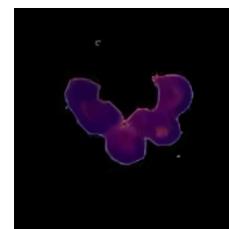


Gambar 4. Citra Hasil Akuisisi (a) dan Citra Hasil Konversi (b)

Gambar di atas memperlihatkan perbedaan citra hasil akuisisi yang masih RGB dengan citra hasil konversi yang sudah menjadi L*a*b. Citra L*a*b selanjutnya digunakan pada langkah segmentasi.

3.3. Segmentasi Citra

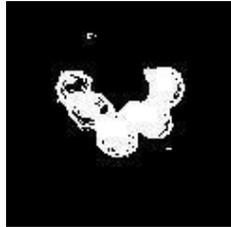
Pada tahap segmentasi proses pertama yang dilakukan adalah clustering. Penelitian ini menggunakan K-means clustering untuk mengelompokkan citra kedalam 3 klaster berdasarkan kedekatannya. Masing-masing klaster akan memberikan hasil yang berbeda, fokus utama penelitian ini adalah klaster inti sel karena akan digunakan pada proses selanjutnya. Hal ini bertujuan untuk mendapatkan inti sel dengan lebih sempurna sebelum diubah menjadi citra biner.



Gambar 5. Citra Hasil Clustering

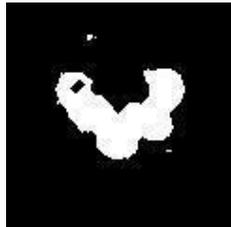
Gambar 5 memperlihatkan citra hasil clustering berfokus pada inti leukosit. Hal tersebut dikarenakan inti sel merupakan ciri khas dari setiap jenis leukosit. citra hasil clustering kemudian diubah menjadi citra biner dengan memperhatikan nilai ambang batas dari objek inti sel dengan latar belakang.

Binerisasi terhadap citra hasil clustering dilakukan dengan menggunakan Otsu thresholding. Otsu thresholding akan membagi citra menjadi dua kelas.



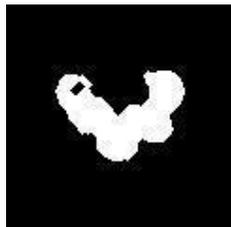
Gambar 6. Citra Hasil Binerisasi

Gambar 6 memperlihatkan bahwa citra hasil binarisasi terhadap hasil clustering masih memiliki bagian yang berlubang. Padahal seharusnya bagian tersebut tersegmentasi menjadi bagian dari objek inti leukosit. Untuk mengisi bagian yang berlubang tersebut maka dilakukan operasi morfologi closing.



Gambar 7. Citra Hasil Operasi Morfologi Closing

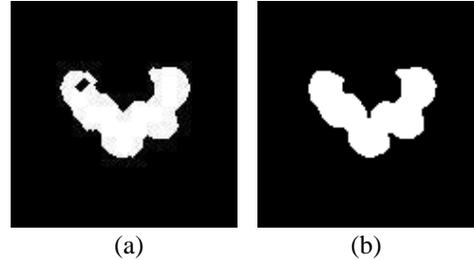
Citra hasil binerisasi pada gambar 6 yang semula masih memiliki banyak lubang sekarang tersegmentasi dengan lebih baik setelah melalui tahap operasi morfologi closing. Hasil segmentasi hampir sempurna, hanya saja masih ada bagian dari artefak atau noise yang juga ikut tersegmentasi, sehingga perlu dilakukan operasi morfologi opening untuk menghapus noise yang tersegmentasi.



Gambar 8. Citra Hasil Operasi Morfologi Opening

Dapat dilihat gambar 8, citra hasil operasi morfologi opening dapat menghilangkan noise yang awalnya ikut

tersegmentasi. Sehingga hasil akhir dari tahap segmentasi semakin mendekati citra ground-truth yang akan digunakan sebagai citra pembandingan dalam evaluasi segmentasi. Adapun perbandingan citra hasil segmentasi dengan citra ground-truth.



Gambar 9. Citra Hasil Segmentasi Akhir (a) dan Citra Ground-truth (b)

3.4. Evaluasi Segmentasi

Hasil segmentasi dievaluasi dengan menggunakan parameter pengukuran PPV dan NPV terhadap piksel citra yang saling bersesuaian antara citra hasil segmentasi akhir dengan citra *ground-truth*. Untuk memperoleh hasil PPV dan NPV maka dilakukan perhitungan *confusion matrix* terhadap perbandingan tersebut. Sehingga diperoleh nilai TP, TN, FP dan FN. Tabel 1 menunjukkan hasil *confusion matrix* dari metode segmentasi *k-means clustering*.

Tabel 1. Hasil *Confusion Matrix k-means*

	TP	TN	FP	FN
Citra 1	49	11951	12	1026
Citra 2	85	8224	18	3970
Citra 3	94	8404	0	4363
Citra 4	80	9431	6	3071
:	:	:	:	:
Citra 299	87	10738	5	2015
Citra 300	90	10337	10	2128

Selanjutnya hasil *confusion matrix* dari penelitian ini dengan mengkombinasikan *k-means clustering* dan *otsu thresholding* ditunjukkan pada Tabel 2.

Tabel 2. Hasil *Confusion Matrix k-means clustering dan Otsu thresholding*

	TP	TN	FP	FN
Citra 1	21	11960	26	453
Citra 2	61	8330	3	3592
Citra 3	43	8404	0	3715
Citra 4	43	9462	5	2586
:	:	:	:	:
Citra 299	47	10748	6	1731
Citra 300	52	10415	5	1904

Berdasarkan nilai *confusion matrix* yang telah diperoleh, kemudian diperoleh nilai PPV dan NPV yang dihitung dengan menggunakan persamaan 10 dan 11. Tabel 3 berikut ini menunjukkan hasil perhitungan PPV dan NPV segmentasi dengan menggunakan *k-means clustering*

berdasarkan *confusion matrix* yang sudah diperoleh sebelumnya seperti pada tabel 1.

Tabel 3. Hasil Evaluasi Segmentasi *k-means clustering*

	PPV	NPV
Citra 1	80,33%	92,09%
Citra 2	82,52%	67,44%
Citra 3	100,00%	65,83%
:	:	:
:	:	:
Citra 299	94,57%	84,20%
Citra 300	90,00%	82,93%
Rata-rata	84,22%	80,91%

Tabel 3 memperlihatkan bahwa rata-rata PPV segmentasi dengan *k-means clustering* adalah 84,22 % sedangkan rerata NPVnya diperoleh sebesar 80,91. Selanjutnya tabel 4 berikut ini menunjukkan hasil perhitungan PPV dan NPV segmentasi kombinasi *k-means clustering* dan *otsu thresholding*.

Tabel 4. Hasil Evaluasi Segmentasi *k-means clustering* dan *Otsu thresholding*

	PPV	NPV
Citra 1	44,68%	96,35%
Citra 2	95,31%	69,87%
Citra 3	100,00%	69,35%
:	:	:
:	:	:
Citra 299	88,68%	86,13%
Citra 300	91,23%	84,54%
Rata-rata	86,33%	81,95%

Tabel 4 memperlihatkan bahwa rata-rata PPV segmentasi dengan kombinasi *k-means clustering* dan *otsu thresholding* adalah 86,33 % sedangkan rerata NPVnya diperoleh sebesar 81,95%. Berdasarkan rerata kedua metode tersebut, dapat dilihat bahwa hasil yang diperoleh masih dalam kisaran persentase yang saling berdekatan. Sehingga perlu dilakukan uji signifikansi untuk mengetahui apakah ada perbedaan terhadap hasil dari kedua metode tersebut dengan menggunakan uji t-test.

Hasil uji signifikansi perbedaan nilai PPV antara dua metode yang diusulkan menghasilkan nilai Sig.(2 tailed) sebesar $0,032 < 0,05$. Hasil tersebut menjelaskan bahwa penggunaan *otsu thresholding* dan morfologi dapat meningkatkan PPV secara signifikan. Selanjutnya, hasil uji signifikansi perbedaan nilai NPV antara dua metode yang diusulkan memperoleh nilai Sig.(2 tailed) sebesar $0,089 > 0,05$. Hasil tersebut menjelaskan bahwa penggunaan *otsu thresholding* dan morfologi tidak meningkatkan nilai NPV secara signifikan.

4. Kesimpulan

Penelitian ini telah berhasil menerapkan algoritma *k-means clustering* untuk memisahkan inti leukosit dengan bagian lainnya yang disempurnakan dengan penggunaan

otsu thresholding dan operasi morfologi. Hasil yang didapatkan bahwa penggunaan *otsu thresholding* dapat meningkatkan performa yang diukur dengan parameter PPV secara signifikan, sedangkan dari segi parameter NPV mengalami peningkatan tetapi tidak signifikan. Pada penelitian ini memiliki kelemahan yaitu kesulitan dalam menentukan kategori cluster, sehingga *cluster* untuk objek inti sel berubah-ubah. Pada penelitian selanjutnya diharapkan dapat menyelesaikan permasalahan pada penelitian ini.

Ucapan Terimakasih

Peneliti mengucapkan terimakasih kepada DIREKTORAT RISET DAN PENGABDIAN MASYARAKAT sebagai pemberi dana, sehingga penelitian ini dapat berjalan dengan lancar.

Daftar Rujukan

- [1] A. R. Andrade, L. H. S. Vogado, R. de M. S. Veras, R. R. V. Silva, F. H. D. Araujo, and F. N. S. Medeiros, "Recent computational methods for white blood cell nuclei segmentation: A comparative study," *Comput. Methods Programs Biomed.*, vol. 173, no. 2019, pp. 1–14, 2019.
- [2] F. Xing and L. Yang, "Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review," *IEEE Rev. Biomed. Eng.*, vol. 9, pp. 234–263, 2016.
- [3] H. Kutlu, E. Avci, and F. Özyurt, "White blood cells detection and classification based on regional convolutional neural networks," *Med. Hypotheses*, vol. 135, no. 2020, pp. 1–11, 2019.
- [4] S. Tantikitti, S. Tumswadi, and W. Premchaiswadi, "Image Processing for Detection of Dengue Virus Based on WBC Classification and Decision Tree," in *International Conference on ICT and Knowledge Engineering*, 2015, pp. 84–89.
- [5] R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, "Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images," *Biocybern. Biomed. Eng.*, vol. 39, no. 2, pp. 382–392, 2019.
- [6] B. Caraka, B. A. A. Sumbodo, and I. Candradewi, "Klasifikasi Sel Darah Putih Menggunakan Metode Support Vector Machine (SVM) Berbasis Pengolahan Citra Digital," *IJEIS (Indonesian J. Electron. Instrum. Syst.)*, vol. 7, no. 1, pp. 25–36, 2017.
- [7] H. Kong, M. Gurcan, and K. Belkacem-Boussaid, "Partitioning Histopathological Images: An Integrated Framework for Supervised Color-texture Segmentation and Cell Splitting," *IEEE Trans. Med. Imaging*, vol. 30, no. 9, pp. 1661–1677, 2011.
- [8] Laddha, S., and S. Daware, "Analysis of White Blood Cell Segmentation Techniques and Classification Using Deep Convolutional Neural Network for Leukemia Detection," *Helix*, vol. 8, no. 6, pp. 4519–4524, 2018.
- [9] X. Hou *et al.*, "A BandMax and spectral angle mapper based algorithm for white blood cell segmentation," in *Ninth International Conference on Digital Image Processing (ICDIP 2017)*, 2017, vol. 10420, no. 104202B, pp. 104202B-1-104202B-5.
- [10] C. Di Ruberto, A. Loddo, and L. Putzu, "A leucocytes count system from blood smear images," *Mach. Vis. Appl.*, vol. 27, no. 8, pp. 1151–1160, 2016.
- [11] Y. Song, W. Cai, H. Huang, Y. Wang, D. D. Feng, and M. Chen, "Region-based progressive localization of cell nuclei in microscopic images with data adaptive modeling," *BMC Bioinformatics*, vol. 14, no. 1, p. 173, 2013.

- [12]R. R. Waliyansyah, "Identifikasi Jenis Biji Kedelai (Glycine Max L) Menggunakan Gray Identification of Glycine Max L Seeds Using Gray Level Coocurance Matrix (Glem) and K-Means Clustering," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 1, pp. 17–26, 2020.
- [13]H. Kong, M. Gurcan, and K. Belkacem-Boussaid, "Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting," *IEEE Trans. Med. Imaging*, vol. 30, no. 9, pp. 1661–1677, Sep. 2011.
- [14]O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234–241.
- [15]C., Zhang, X., Xiao, X., Li, Chen, J., Ying, W., Zhen, J., Chang, C., Zheng, Z., Liu, "White blood cell segmentation by color-space-based k-means clustering," *Sensors (Switzerland)*, vol. 14, no. 9, pp. 16128–16147, 2014.
- [16]N. Theera-Umpon, "White blood cell segmentation and classification in microscopic bone marrow images," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3614 LNAI, pp. 787–796, 2005.
- [17]S. Arslan, E. Ozyurek, and C. Gunduz-Demir, "A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images," *Cytometry*, vol. 85, pp. 480–490, 2014.
- [18]L. B. Dorini, R. Minetto, and N. J. Leite, "White blood cell segmentation using morphological operators and scale-space analysis," in *XX Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI 2007)*, 2007, pp. 294–304.
- [19]M. Farhan, O. Yli-Harja, and A. Niemistö, "A novel method for splitting clumps of convex objects incorporating image intensity and using rectangular window-based concavity point-pair search," *Pattern Recognit.*, vol. 46, no. 3, pp. 741–751, 2013.
- [20]M. G. Abdul-Haleem, "White blood cells nuclei localization using modified K-means clustering algorithm and seed filling technique," *Iraqi J. Sci.*, vol. 60, no. 4, pp. 875–890, 2019.
- [21]A. Harto and C. Fatichah, "Segmentasi Dan Pemisahan Sel Darah Putih Bersentuhan Menggunakan K-Means Dan Hierarchical Clustering Analysis Pada Citra Leukemia Myeloid Akut," *JUTI J. Ilm. Teknol. Inf.*, vol. 15, no. 2, p. 162, 2017.
- [22]I. Herawati, Faridah, B. Achmad, and R. J. Yanti, "The effect of contrast enhancement method for k-means clustering segmentation of white blood cell cytoplasm image," *J. Eng. Sci. Technol.*, vol. 15, no. 1, pp. 227–248, 2020.
- [23]N. Thomas and V. Sreejith, "A Review on White Blood Cells Segmentation," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 396, no. 2018, pp. 1–5, 2018.
- [24]Sutrisno, A. A. Supianto, and I. Cholissodin, "Implementasi Teknik Watershed Dan Morfologi Pada Citra Satelit Untuk Segmentasi Area Universitas Brawijaya," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 1, no. 1, p. 5, 2014.
- [25]M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.
- [26]M. Fatourehchi, R. K. Ward, S. G. Mason, J. Huggins, A. Schlögl, and G. E. Birch, "Comparison of evaluation metrics in classification applications with imbalanced datasets," in *Proceedings - 7th International Conference on Machine Learning and Applications, ICMLA 2008*, 2008, pp. 777–782