

Klasifikasi Sel Darah Putih Berdasarkan Ciri Warna dan Bentuk dengan *Metode K-Nearest Neighbor (K-NN)*

Mizan Nur Khasanah^{*1}, Agus Harjoko², Ika Candradewi³

¹Prodi Elektronika dan Instrumentasi Jurusan Ilmu Komputer dan Elektronika, FMIPA UGM

^{2,3}Jurusan Ilmu Komputer dan Elektronika, FMIPA UGM

e-mail: ^{*1}miza.nk19@gmail.com, ²aharjoko@ugm.ac.id ³ika-candra@mail.ugm.ac.id

Abstrak

Prosedur tradisional klasifikasi sel darah menggunakan mikroskop di laboratorium hematologi dilakukan untuk memperoleh informasi jenis sel darah. Telah menjadi landasan di laboratorium hematologi untuk mendiagnosis dan memantau gangguan hematologi. Namun, prosedur manual melalui serangkaian uji laboratorium dapat memakan waktu cukup lama. Oleh karena itu penelitian ini ditujukan khusus untuk dapat membantu dalam proses tahap awal klasifikasi jenis sel darah putih secara otomatis di bidang medis.

Upaya untuk mengatasi lamanya waktu dan untuk keperluan diagnosis awal dapat menggunakan teknik pengolahan citra berdasarkan morfologi sel darah. Penelitian ini bertujuan untuk mengklasifikasi sel darah putih berdasarkan morfologi sel dengan k-nearest neighbor (knn). Algoritma pengolahan citra yang digunakan adalah hough circle, thresholding, ekstraksi ciri. kemudian untuk proses klasifikasi digunakan metode k-nearest neighbor (knn).

Pada proses pengujian digunakan 100 citra untuk di ketahui jenisnya. Hasil pengujian segmentasi menunjukkan akurasi sebesar 78 % dan pengujian klasifikasi sebesar 64%.

Kata kunci— Sel darah putih, hough circle, ekstraksi ciri, k-nearest neighbor (k-nn)

Abstract

The traditional procedure of classification of blood cells using a microscope in the laboratory of hematology to obtain information types of blood cells. It has become a cornerstone in the laboratory of hematology to diagnose and monitor hematologic disorders. However, the manual procedure through a series of labory test can take a while. Thresfore, this research can be helpful in the early stages of the classification of white blood cells automatically in the medical field.

Efforts to overcome the length of time and for the purposes of early diagnose can use the image processing technique based on morphology of blood cells. This research aims to classify the white blood cells based on cell morphology with the k-nearest neighbor (knn). Image processing algorithms used hough circle, thresholding, feature extraction, then to the process of classification was used the method of k-nearest neighbor (knn).

In the process of testing used 100 images to be aware of its kind. The test results showed segmentation accuracy of 78% and testing the classification of 64%.

Keywords—White blood cell, hough circle, feature extraction, k-nearest neighbor (knn)

1. PENDAHULUAN

Darah di dalam tubuh manusia memiliki fungsi yang sangat penting sebagai alat untuk transportasi oksigen dan zat-zat yang dibutuhkan oleh tubuh. Darah merupakan cairan tubuh yang berwarna merah, warna merah ini merupakan protein pernafasan yang mengandung besi, yang merupakan tempat terikatnya molekul-molekul oksigen yang disebabkan oleh hemoglobin. Dalam darah juga terdapat kandungan seperti air, protein, mineral dan garam. Selain itu darah juga dibedakan menjadi beberapa jenis. Pada masing-masing jenis darah juga memiliki peranan penting dalam tubuh. Jenis-jenis darah manusia yakni sel darah merah, sel darah putih serta kepingan darah [1].

Sel darah putih merupakan salah satu bagian dari susunan sel darah manusia yang memiliki peranan utama dalam hal sistem imunitas atau membunuh kuman dan bibit penyakit yang ikut masuk ke dalam aliran darah manusia. Sel darah putih atau yang juga dapat disebut dengan leukosit. Leukosit dibagi menjadi lima jenis tipe berdasarkan bentuk morfologinya yaitu *basofil*, *eosinofil*, *neutrofil*, *limfosit* dan *monosit* [2]. Masing-masing jenis sel darah putih ini memiliki ciri khas dan fungsi yang berbeda.

Berdasarkan pada jurnal pemeriksaan hematologi klinik 1 oleh Heru Santoso Wahito Nugroho, S.Kep., Ns., M.M.Kes., selama ini proses klasifikasi sel darah putih hanya dilakukan secara manual melalui serangkaian uji laboratorium yang dapat memakan waktu cukup lama. Oleh karena itu penelitian ini ditujukan khusus untuk dapat membantu dalam proses tahap awal klasifikasi jenis sel darah putih secara otomatis di bidang medis.

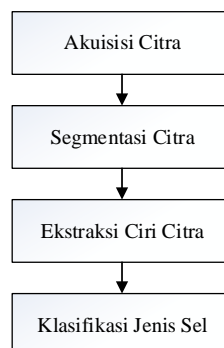
Seiring dengan berkembangnya teknologi, saat ini ada beberapa peneliti yang melakukan berbagai penelitian dalam penerapan metode komputerasi yang digunakan dalam klasifikasi jenis sel darah putih. Beberapa diantaranya adalah identifikasi dan klasifikasi sel darah putih menggunakan segmentasi berbasis warna dan fitur geometris yang dilakukan oleh Hiremath, P.S, dkk pada tahun 2010. Cortes, R pada tahun 2011 melakukan penelitian klasifikasi sel darah putih dengan metode pengenalan pola *euclidean distance*, *k-nearest neighbor* dengan hasil yang cukup baik pada citra yang memiliki kedekatan ciri [3].

Mengacu pada penelitian sebelumnya, maka penulis akan menggabungkan proses segmentasi citra sel darah putih dengan *k-nearest neighbor* dimana proses ekstraksi ciri digunakan sebagai parameter inputan ke dalam proses pengujian klasifikasi sel darah putih.

2. METODE PENELITIAN

2.1 Diskripsi Sistem

Sistem yang dibuat merupakan sistem untuk melakukan klasifikasi sel darah putih dari citra darah. Berikut adalah blok diagram rancangan pembuatan sistem yang ditunjukkan pada Gambar 1.



Gambar 1 Blok diagram rancangan sistem

Proses klasifikasi sel darah putih membutuhkan beberapa tahapan. Secara umum, proses klasifikasi sel darah putih meliputi akuisisi citra, segmentasi citra, ekstraksi fitur, dan klasifikasi dengan metode *k-nearest neighbor* (knn). Sistem menerima masukan berupa citra darah dan dilakukan proses segmentasi yang berguna untuk mendeteksi sel darah putih di dalam citra darah. Hasil dari segmentasi adalah berupa sel darah putih yaitu sebagai objek yang selanjutnya akan di kenai operasi *hough circle*. Citra sel darah putih yang berhasil dideteksi kemudian dilakukan ekstraksi ciri. Hal ini digunakan untuk mendapatkan ciri masing-masing sel darah putih. Metode *k-nearest neighbor* (knn) digunakan untuk melakukan klasifikasi jenis sel darah putih berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut [4]. Citra yang sudah di ekstraksi ciri akan dicocokkan dengan sampel yang sudah diketahui jenis-jenisnya sehingga hasil akhir sistem ini adalah mengetahui jenis sel darah putih yang terdapat pada citra darah.

2.2 Perancangan Akuisisi Citra

Proses awal adalah akuisisi data Citra. Akuisisi citra bertujuan untuk mendapatkan citra digital dari mikroskopik preparat sel darah. Citra digital diperoleh dari *scanning* foto mikroskopis sel darah dalam 24 bit warna dengan penyimpanan format jpg yang didapat dari laboratorium rumah sakit Sardjito. Citra yang dipilih adalah citra 24 bit yang dikenali sebagai citra RGB sehingga citra siap untuk diproses pada tahap selanjutnya.

2.3 Perancangan Segmentasi Citra

Diagram alir rancangan segmentasi citra terdapat pada Gambar 2. Proses ini mendapat masukan citra darah. Proses deteksi sel darah putih ini dilakukan dengan metode *hough circle*. Proses awal deteksi dengan *hough circle* adalah *grayscale*, yaitu mengubah warna citra asli berupa RGB menjadi citra *grayscale* [5]. Hasil dari proses *hough* ini ditandai dengan sel darah dengan lingkaran berwarna merah. Koordinat dari obyek tersebut digunakan untuk mengatur *region of interest* (ROI). Fungsi dari ROI adalah untuk mengatur area dari sel darah yang terdeteksi [6].

Selanjutnya dilakukan operasi untuk menghitung rata-rata warna objek yang terdeteksi. Operasi ini digunakan untuk memverifikasi nilai rata-rata warna objek yang terdeteksi. Jika hasil rata-rata warna dari objek terdeteksi adalah >120 maka variabel *circleF* akan ditambahkan ke dalam *listCircle*. Sehingga dengan menggunakan verifikasi rata-rata warna tersebut hanya citra sel darah putih yang dapat terdeteksi karena warna citra sel darah putih yang cenderung berwarna gelap. Citra sel darah putih yang terdeteksi selanjutnya dilakukan *cropping* sesuai dengan area yang telah ditentukan oleh ROI tersebut. Citra hasil *cropping* tersebut yang akan ditampilkan pada *PictureBox*.

2.4 Perancangan Ekstraksi Ciri

Ekstraksi ciri dilakukan pada citra sel darah putih yang telah berhasil disegmentasi. Tujuan dilakukannya proses ekstraksi ciri ini adalah untuk mendapatkan data numerik ciri masing-masing citra dalam parameter ciri citra, yaitu luas area, tepi area, kebundaran sel, dan rata-rata warna *red*, *green*, dan *blue*.

Diagram alir proses ekstraksi ciri ditunjukkan pada Gambar 3. Masukan pada sistem ini adalah citra hasil *cropping* pada proses segmentasi yang selanjutnya dilakukan operasi *thresholding* yang bertujuan untuk mempermudah perhitungan parameter-parameter ciri sel. Citra hasil *thresholding* berupa citra biner yaitu obyek dengan area warna hitam. Selanjutnya obyek tersebut akan di ekstraksi ciri dengan parameter-parameter yang telah ditentukan.

Berikut merupakan perancangan perhitungan nilai parameter-parameter tersebut:

Luas Area

Luas area citra segmentasi sel darah putih didapatkan dari ukuran luas badan sel atau sitoplasma sel darah putih. Piksel area sitoplasma hasil segmentasi yang berwarna hitam akan

diberi nilai 1. Sedangkan *background* akan bernilai 0. Luas area dapat diperoleh dengan menghitung jumlah piksel bernilai 1 [7].

Tepi Area

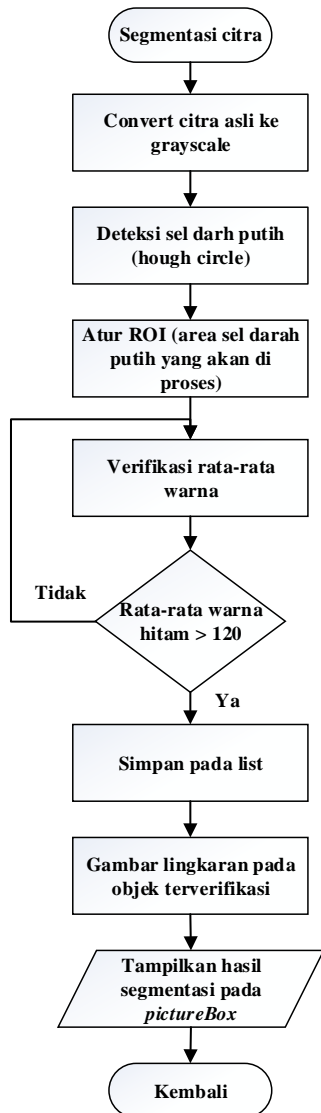
Tepi area atau perimeter merupakan bagian sitoplasma yang paling luar dan bersinggungan dengan *background* citra. Tepi citra dapat diperoleh dengan menghitung piksel terluar dari objek. Dengan demikian akan didapatkan tepi area yang kemudian dapat dihitung jumlah pikselnya. Piksel tepi area berwarna putih akan bernilai 1, sedangkan sisanya akan bernilai 0. Tepi nukleus dapat diperoleh dengan menghitung jumlah piksel bernilai 1 tersebut [7].

Kebundaran

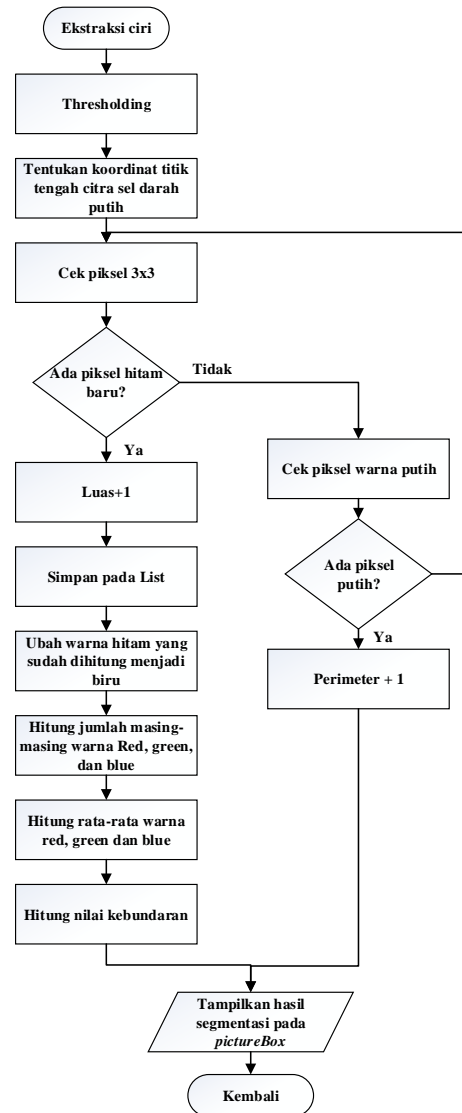
Kebundaran atau tingkat kelengkungan dari sel darah putih dapat dihitung setelah diperoleh nilai luas dan tepi area melalui rumus kebundaran. Nilai dari kebundaran akan berkisar antara 0 sampai 1. Semakin bundar citra sel darah putih, maka akan semakin tinggi nilai kebundarannya atau mendekati 0 [8].

Rata-Rata RGB

Rata-rata warna *red*, *green*, dan *blue* merupakan rata-rata nilai warna pada sel darah putih yang terdeteksi. Nilai rata-rata warna ini didapat dengan menghitung masing-masing warna kemudian dibagi dengan luas area objek tersebut.



Gambar 2 Diagram alir segmentasi citra

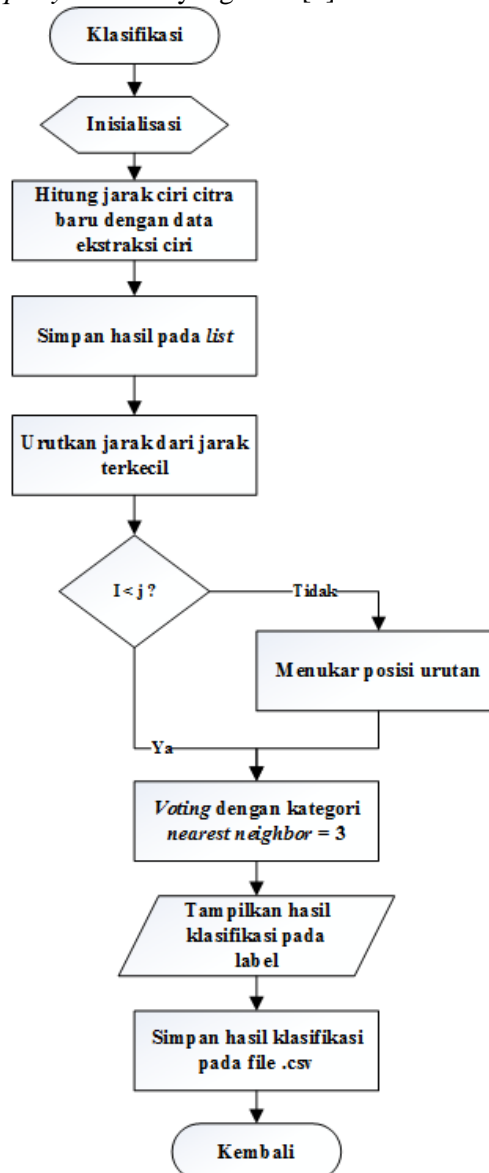


Gambar 3 Diagram alir ekstraksi ciri citra

2.5 Perancangan Klasifikasi

Klasifikasi sel darah putih dilakukan menggunakan metode *k-nearest neighbor* (KNN) yaitu sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Tujuan dari algoritma ini adalah untuk mengklasifikasikan sel darah putih baru berdasarkan atribut dan *training sample* jenis sel darah putih yang disimpan dengan format csv [4]. Diagram alir proses klasifikasi ditunjukkan pada Gambar 4.

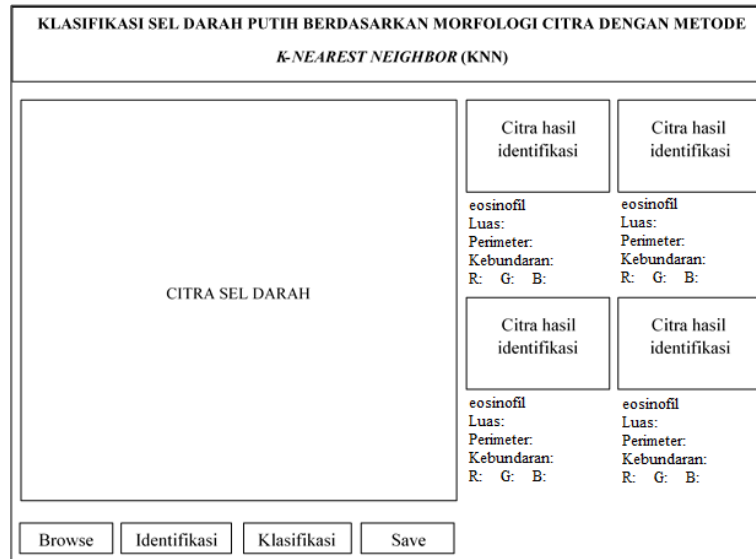
Training sample diproyeksikan ke ruang berdimensi banyak, dimana masing masing dimensi merepresentasikan fitur dari masing masing sel darah putih. *Classifier* tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Diberikan titik *query* (sel darah putih baru), kemudian akan ditemukan 3 objek atau titik *training* yang paling dekat dengan titik *query*. Klasifikasi menggunakan *voting* terbanyak diantara klasifikasi dari 3 obyek. Algoritma *k-nearest neighbor* (KNN) menggunakan klasifikasi ketetanggan sebagai nilai prediksi dari *query instance* yang baru [4].



Gambar 4 Diagram alir klasifikasi citra

2.6 Perancangan Antarmuka Sistem

Antarmuka aplikasi pada penelitian ini digunakan untuk memudahkan peneliti dalam melihat hasil klasifikasi sel darah putih. Antarmuka aplikasi memuat fungsionalitas pemuatan citra, pengidentifikasian citra sel darah putih dan penentuan jenis sel. Rancangan antarmuka aplikasi ditunjukkan pada Gambar 5



Gambar 5 Rancangan antarmuka sistem

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengujian Segmentasi Citra

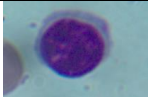


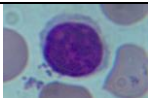
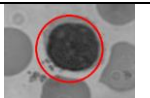

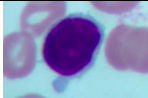


Segmentasi Citra sel darah putih dengan metode *hough circle* dilakukan pada citra sebanyak 128 buah. Metode *hough circle* diimplementasikan pada aplikasi ini karena obyek yang akan dideteksi berbentuk lingkaran. Tujuan dari proses segmentasi ini adalah untuk mendeteksi sel darah putih di dalam suatu citra, menyiapkan citra ke kondisi untuk siap diekstraksi ciri dan kemudian memperoleh data numerik dari hasil ekstraksi ciri citra tersebut. Pencarian obyek sel darah putih sangat ditentukan oleh parameter-parameter dari metode *hough circle*. Penentuan parameter *hough circle* yang tidak sesuai dalam proses segmentasi ini dapat menyebabkan kesalahan dalam mendeteksi sel darah putih. Beberapa kesalahan yang terjadi adalah adanya sel darah merah dan sel darah rusak yang terdeteksi. Beberapa kesalahan tersebut merupakan kelemahan dari metode *hough circle*, karena metode ini ditentukan oleh parameter-parameter yang tidak dapat dipastikan untuk mendapatkan hasil terbaik dari satu citra dengan citra yang lainnya. Sehingga untuk beberapa citra harus dilakukan pengaturan ulang nilai parameternya dengan cara *trial and error* untuk dapat mendeteksi sel darah putih dengan baik. Pada penelitian ini telah dilakukan proses *trial and error* dalam mencari nilai parameter-parameter *hough circle* yang dapat mewakili proses deteksi obyek seluruh citra yang diproses. Hasil terbaik yang dapat mendeteksi sel darah putih adalah nilai *accumulator* = 133; *canny* = 246; *maxRadius* = 148; *minRadius* = 60; *minDistance* = 90; dan *resolution* = 30.

Secara umum, sistem aplikasi dapat digunakan untuk mengidentifikasi sel darah putih yang ada di dalam citra dengan menggunakan parameter-parameter tersebut. Hasil segmentasi citra adalah representasi dari sel darah putih dalam format *grayscale*.

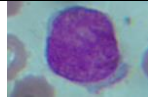


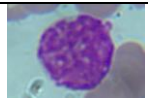
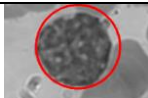
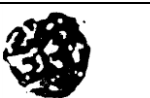
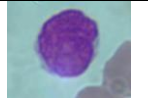


Contoh hasil segmentasi yang berhasil ditunjukkan pada Tabel 1. Pada tabel tersebut terdiri dari citra asli, citra hasil deteksi dan citra hasil segmentasi. Citra hasil deteksi dengan lingkaran warna hitam adalah hasil dari proses *hough circle*. Sedangkan hasil segmentasi merupakan proses *thresholding* yang selanjutnya dilakukan perhitungan ciri citra.

Contoh hasil segmentasi gagal ditunjukkan pada Tabel 2. Pada tabel tersebut terlihat bahwa hasil segmentasi berbeda dengan hasil segmentasi yang berhasil. Pada hasil segmentasi yang berhasil, obyek berwarna biru karena nilai piksel yang sudah ditandai akan dihitung kemudian diganti warna biru untuk menandai bahwa piksel tersebut sudah dihitung. Sedangkan pada segmentasi yang gagal obyek tetap berwarna hitam karena perhitungan ciri dimulai dari titik tengah obyek, sedangkan pada citra tersebut titik tengah berwarna putih, sehingga tidak dapat dilakukan perhitungan maka obyek tidak berganti dengan warna biru.

Tabel 1 Contoh hasil segmentasi berhasil

Citra Asli	Hasil Deteksi	Hasil Segmentasi
		
		
		

Tabel 2 Contoh hasil segmentasi gagal

Citra Asli	Hasil Deteksi	Hasil Segmentasi
		
		
		

3.2 Hasil Pengujian Ekstraksi Ciri Citra

Ekstraksi ciri citra digunakan untuk mencari parameter yang digunakan untuk mencocokkan dengan data ciri sampel sel darah yang sudah di ekstraksi yaitu luas area, tepi area, kebundaran, rata-rata warna *red*, *green* dan *blue*. Luas area dihitung berdasarkan jumlah piksel yang berada di dalam hasil segmentasi *nukleus* dan *sitoplasma*. Tepi area sel atau *perimeter* juga dihitung dalam satuan piksel dari tepi *sitoplasma*. Kebundaran merupakan rasio yang didapat dari perhitungan antara luas dan tepi area dalam rentang 0 sampai 1. Semakin tinggi rasio kebundaran maka semakin bundar bentuk sel. Rata-rata warna *red*, *green* dan *blue* merupakan nilai masing-masing warna sel dibagi dengan luas area.

Adapun hasil rata-rata yang didapatkan dari ekstraksi ciri ditunjukkan pada Tabel 3 dengan A: rata-rata; B: nilai tertinggi; C: nilai terendah. Hasil ekstraksi ciri pada Tabel 3 Menunjukkan bahwa *neutrofil* dan *basofil* memiliki kedekatan ciri. Kedua jenis sel tersebut menghasilkan rata-rata nilai dari enam parameter dengan perbedaan yang tidak signifikan. Kemudian sel jenis *eosinofil* dan *limfoblast* juga memiliki kedekatan ciri, terlihat dari rata-rata luas dan kebundaran keduanya. Pada sel *limfosit* terlihat bahwa nilai rata-rata kebundaran yang dihasilkan adalah yang tertinggi daripada keempat sel lainnya. Meskipun demikian, rata-rata luas dan *perimeter* yang dimilikinya adalah yang terendah. Sedangkan pada *monosit* nampak bahwa rata-rata luas area dan *perimeter* yang dihasilkan adalah yang tertinggi. Kemudian jika dilihat dari rata-rata warna *red*, *green*, dan *blue*, sel *neutrofil*, *monosit* dan *limfosit* memiliki rata-rata warna biru yang dominan dibandingkan dengan sel darah lainnya.

Jika dilihat berdasarkan kondisi nyata di dalam tubuh manusia, hasil ekstraksi pada beberapa sel tersebut cukup mampu merepresentasikan bentuk identik sel. Sebagai contoh, sel terbesar di dalam tubuh manusi adalah *monosit*. Hasil ekstraksi ciri juga telah menunjukkan bahwa rata-rata terbesar pada luas area yang merepresentasikan ukuran sel dimiliki oleh citra *monosit*.

Tabel 3 Hasil rata-rata ekstraksi ciri

Jenis		Neutrofil	Eosinofil	Basofil	Limfosit	Monosit	Limfoblast
Luas	A	32806.38	217343.4	28972.33	13555.66	33401.11	20496.86
	B	38394	267770	35741	15608	50495	27939
	C	30236	152216	20165	10545	17585	15395
Perimeter	A	1086.077	2670.8	1068.333	550.5862	1380.222	718.6094
	B	1729	3679	1847	650	2560	1001
	C	818	2169	648	464	905	581
Kebundaran	A	0.039565	0.04014	0.039417	0.057207	0.029678	0.052122
	B	0.0596	0.0459	0.0612	0.0628	0.0594	0.0673
	C	0.0141	0.0252	0.0133	0.0471	0.0095	0.0317
Rata-rata Red	A	116.6758	85.08	111.39	120.0859	125.23	83.42906
	B	188.82	106.11	141.82	157.76	164.97	170.26
	C	88.97	56.89	88.37	83.42	99.45	46.01
Rata-rata green	A	81.67808	32.134	74.29833	51.6069	73.60556	29.88625
	B	134.23	46.21	116.94	101.82	130.55	94.88
	C	20.25	14.73	39.89	29.07	37.3	1.17
Rata-rata blue	A	154.3692	116.95	144.2933	167.9459	159.5489	141.5528
	B	208.79	139.33	157.84	200.13	191.57	198.25
	C	139.28	94.38	130.19	125.59	148.12	132.02

3.3 Hasil Pengujian Klasifikasi Citra Sel Darah Putih

Tujuan dari algoritma k-NN ini adalah untuk mengklasifikasi citra baru berdasarkan data ekstraksi ciri yang telah diketahui jenisnya. Dimana hasil dari sampel uji citra baru diklasifikasikan berdasarkan mayoritas dari kategori pada NN. Pengujian terhadap metode KNN dilakukan menggunakan parameter $k=3$. Parameter K merupakan jumlah tetangga paling dekat dalam hal ini digunakan untuk menentukan hasil klasifikasi. Jarak yang digunakan adalah jarak *Euclidean Distance*. *Jarak Euclidean* adalah jarak yang paling umum digunakan pada data numerik. Pada pengujian ini dilakukan dengan 100 citra darah baru yang belum diketahui jenisnya. Klasifikasi akan dibedakan menjadi 6 jenis yaitu *basofil*, *eosinofil*, *neutrofil*, *limfosit*, *monosit* dan *limfoblast*. Hasil pengujian dengan metode *k-nearest neighbor* ditunjukkan pada Tabel 4.

Dapat dilihat bahwa hasil klasifikasi dengan metode KNN menunjukkan bahwa dari 100 citra baru, sistem dapat mendeteksi enam jenis sel darah. Jenis sel darah putih yang dominan terdeteksi adalah sel *limfosit* dan *limfoblast*. Sedangkan sel darah *monosit*, *neutrofil*, *basofil*, dan *eosinofil* hanya terdeteksi pada beberapa citra saja. Misalnya sel jenis *monosit* hanya terdeteksi pada 8 citra, sel jenis *neutrofil* terdeteksi pada 2 citra, sel jenis *basofil* terdeteksi pada 3 citra dan sel jenis *eosinofil* terdeteksi pada 7 citra. Pada dasarnya sel jenis *Basofil* sangat sukar ditemukan dalam apusan sel darah karena jumlahnya kurang dari 1% sel darah putih [9].

Ada beberapa faktor yang mempengaruhi keberhasilan algoritma KNN ini antara lain adalah jumlah data masing-masing jenis sel darah. Pada penelitian ini jumlah data *basofil*, *eosinofil*, *neutrofil* dan *monosit* lebih sedikit dibandingkan dengan data *limfosit* dan *limfoblast*. Perbandingan jumlah data yang sangat jauh mungkin hasil klasifikasi yang kurang akurat, karena kemungkinan sistem mendeteksi jarak yang terdekat adalah pada data yang lebih banyak. Faktor lain yang mempengaruhi keberhasilan algoritma ini adalah proses segmentasi citra. Segmentasi citra yang gagal terutama pada tahap *thresholding* akan mempengaruhi keakuratan ekstraksi ciri. Jika proses *thresholding* tidak sempurna, maka hasil ekstraksi ciri akan berkurang dan atau bertambah lebih besar sehingga nilai ciri yang akan dicocokkan akan menjadi tidak *valid*. Sehingga keberhasilan dalam proses segmentasi sangat berpengaruh pada hasil klasifikasi ini.

Tabel 4 Hasil pengujian dengan metode *k-nearest neighbor*

No	Nama File	Luas	Perimeter	Kebundaran	Rata-rata Re	Rata-rata Gree	Rata-rata Blu	Klasifikasi
1	3 citra 1	15403	811	0.0298	51.67	11.89	120.9	limfosit
2	5 citra 1	13709	577	0.0525	75.75	36.75	121.97	limfosit
3	8 citra 1	20584	754	0.0461	88.3	36.66	131.15	limfoblast
4	9 citra 1	24020	1059	0.0273	90.71	39.1	141.26	limfoblast
5	10 citra 1	23755	819	0.0451	85.72	41.25	135.8	limfoblast
6	11 citra 1	14693	565	0.0586	80.01	43.54	127.87	limfosit
7	14 citra 1	9564	357	0.0956	76.85	37.38	128.94	limfosit
8	15 citra 1	11476	543	0.0496	70.97	35.6	122.66	limfosit
9	16 citra 1	10509	448	0.0667	75.46	30.89	130.11	limfosit
10	17 citra 1	22822	659	0.0669	87.63	34.98	135.07	limfoblast
11	19 citra 1	25091	1287	0.0193	90.3	49.79	135.26	limfoblast
12	20 citra 1	10870	488	0.0581	70.52	31.16	125.76	limfosit
13	25 citra 1	10724	482	0.0588	75.82	47.3	128.53	limfosit
14	30 citra 1	24909	1867	0.0091	100.46	51.44	142.46	limfoblast
15	31 citra 1	14484	556	0.0597	79.01	35.52	127.73	limfosit
16	33 citra 1	21978	1084	0.0238	99.32	42.98	144.73	limfoblast
17	34 citra 1	29073	1689	0.013	97.95	53.09	144.39	neutrofil
18	36 citra 1	23439	1526	0.0128	86.75	49.13	132.62	limfoblast
19	39 citra 1	17876	706	0.0457	93.63	38.28	139.02	limfoblast
20	43 citra 1	9215	455	0.0567	65.64	45.68	123.65	limfosit
21	44 citra 1	8445	435	0.0569	65.93	48.28	123.45	limfosit
22	48 citra 1	21515	991	0.0279	92.56	59.47	139.88	limfoblast
23	54-55 citra 1	11329	530	0.0514	75.83	30.85	124.61	limfosit
24	56-57 citra 1	14369	589	0.0528	75.55	35.05	123.96	limfosit
25	60 citra 1	22114	1324	0.0161	84.34	54.88	129.21	limfoblast
26	61 citra 1	30891	1312	0.0229	90.53	50.99	137.27	neutrofil
27	66 citra 1	10888	486	0.0587	71.14	42.06	122.37	limfosit
28	67 citra 1	8699	476	0.0489	67.13	39.04	120.6	limfosit
29	68 citra 1	13742	601	0.0485	85.61	42.73	135.32	limfosit
30	69 citra 1	10972	324	0.1331	70.91	30.63	122.84	limfosit
31	71 citra 1	10769	432	0.0735	72.97	28.78	126.55	limfosit
32	72 citra 1	12317	576	0.0473	72.39	26.85	126.84	limfosit
33	76 citra 1	11789	450	0.0742	73.69	33.33	124.7	limfosit
34	79 citra 1	23645	884	0.0385	83.24	35.01	135.19	limfoblast
35	80 citra 1	24235	1002	0.0307	84.54	44.56	133.9	limfoblast
36	81 citra 1	15448	1117	0.0158	77.9	46.06	132.53	limfosit
37	83 citra 1	21255	1059	0.0241	82.85	30.56	135.32	limfoblast
38	85 citra 1	14312	589	0.0526	76.24	33.66	126.84	limfosit
39	97 citra 1	15144	1269	0.012	75.57	35.8	130.84	limfosit
40	101 citra 1	14150	588	0.0521	71.23	35.87	125.03	limfosit
41	10 citra 1	16243	841	0.0293	95.96	41.17	136.59	limfoblast
42	11 citra 1	14691	659	0.0431	95.23	43.83	138.05	limfosit
43	17 citra 1	24846	1989	0.008	94.27	45.07	138.23	limfoblast
44	99 citra 1	17608	1071	0.0196	99.35	45.69	137.45	limfoblast
45	19 citra 1	12445	488	0.0666	89.22	36.18	134.16	limfosit
46	22 citra 1	14492	647	0.0441	88.71	34.11	129.2	limfosit
47	29 citra 1	26281	1140	0.0258	84.43	33.23	127.72	eosinofil
48	30 citra 1	15502	780	0.0325	96.56	41.3	138.6	limfosit
49	37 citra 1	16782	842	0.0302	96.95	38.97	140.7	limfoblast
50	46 citra 1	25672	713	0.0643	102.37	35.86	135.98	eosinofil

No	Nama File	Luas	Perimeter	Kebundaran	Rata-rata Re	Rata-rata Gree	Rata-rata Blu	Klasifikasi
51	47 citra 1	20574	370	0.1914	98.82	38.74	138.73	limfoblast
52	48 citra 1	17059	637	0.0536	100.16	44.92	142.65	limfoblast
53	88 citra 1	20644	1150	0.0199	94.64	44.85	133.57	limfoblast
54	2 citra 1	26635	2843	0.0042	117.32	35.87	165.53	basofil
55	3 citra 1	23684	1051	0.0273	114.74	39.75	159.43	limfoblast
56	4 citra 1	28448	2206	0.0074	115.09	32.56	162.92	basofil
57	36 citra 1	21596	1520	0.0119	107.72	57.47	142.8	limfoblast
58	38 citra 1	14919	1955	0.005	116.06	57.5	150.8	limfosit
59	43 citra 1	16383	2910	0.0025	114.9	58.13	150.19	limfoblast
60	47 citra 1	22167	1671	0.0101	107.73	52.88	146.48	limfoblast
61	48 citra 1	12750	2010	0.004	115.83	59.11	151	limfosit
62	49 citra 1	18171	2046	0.0055	113.07	56.67	146.99	limfoblast
63	53 citra 1	14920	2347	0.0035	114.9	58.27	152.61	limfosit
64	54 citra 1	16666	1327	0.0121	108.29	55.62	147.08	limfoblast
65	58 citra 1	17635	776	0.0373	103.41	46.26	139.26	limfoblast
66	65 citra 1	13971	819	0.0265	105.3	44.36	144.66	limfosit
67	66 citra 1	14432	925	0.0215	103.86	45.82	143.51	limfosit
68	67 citra 1	16699	918	0.0252	111.21	50.94	145.38	limfoblast
69	68 citra 1	16572	981	0.0219	103.42	50.34	138.88	limfoblast
70	71 citra 1	17059	1656	0.0079	108.83	58.49	146.43	limfoblast
71	76 citra 1	22280	2602	0.0042	109.55	61.37	143.08	limfoblast
72	77 citra 1	19970	2917	0.003	113.01	60.83	144.35	limfoblast
73	80 citra 1	16835	3918	0.0014	114.49	62.06	150.47	limfoblast
74	81 citra 1	19261	3034	0.0027	111.92	60.29	148.84	limfoblast
75	83 citra 1	13126	1536	0.0071	110.31	56.75	143.91	limfosit
76	86 citra 1	18432	520	0.0868	105.08	48.91	142.88	limfoblast
77	87 citra 1	28358	3324	0.0033	118.2	52.13	158.72	basofil
78	98 citra 1	16340	2485	0.0034	114.22	53.67	155.79	limfoblast
79	106 citra 1	13682	778	0.0288	103.41	45.66	142.46	limfosit
80	107 citra 1	13771	1298	0.0104	107.03	54.82	148.12	limfosit
81	33 citra 1	19590	697	0.0514	61.03	5.43	133.16	limfoblast
82	43 citra 1	16008	613	0.0543	87.53	45.61	131.44	limfoblast
83	73 citra 1	12758	266	0.2297	95.6	35.67	142.89	limfosit
84	74 citra 1	22579	1625	0.0109	107.83	47.34	153.12	limfoblast
85	75 citra 1	17650	860	0.0304	96.91	38.86	140.14	limfoblast
86	76 citra 1	10847	414	0.0806	105.26	45.35	146.25	limfosit
87	137 citra 1	91751	1115	0.094	77.64	4.44	121.98	monosit
88	139 citra 1	115320	1040	0.1358	88	3.15	129.89	eosinofil
89	141 citra 1	126307	1808	0.0492	80.84	4.83	129.28	eosinofil
90	148 citra 1	108130	1473	0.0635	70.7	4.04	124.05	monosit
91	152 citra 1	125575	1582	0.0639	73.22	3.62	117.19	eosinofil
92	158 citra 1	125352	1530	0.0682	88.51	3.92	131.32	eosinofil
93	159 citra 1	100101	1043	0.1172	86.14	2.77	127.52	monosit
94	118 citra 1	99834	1354	0.0694	85.86	5.28	120.32	monosit
95	114 citra 1	15825	214	0.4402	56.5	3.61	107.75	limfosit
96	108 citra 1	97967	1758	0.0404	70.32	7.71	109.96	monosit
97	98 citra 1	95885	1474	0.0562	67.5	4.3	109.02	monosit
98	97 citra 1	96365	999	0.123	86	3.4	116.61	monosit
99	94 citra 1	113410	2564	0.022	89.26	6.6	120.22	eosinofil
100	72 citra 1	69132	1310	0.0513	54.61	6.82	104.46	monosit

Faktor yang lebih berpengaruh pada algoritma KNN ini adalah parameter-parameter ciri yang digunakan. Pada penelitian ini hanya menggunakan 6 parameter ciri yaitu luas area, *perimeter*, *kebundaran* dan rata-rata warna *red*, *green* dan *blue*. Dari hasil pengujian ekstraksi ciri diatas dapat dilihat bahwa ada beberapa sel darah yang mempunyai kedekatan ciri yaitu *basofil* dengan *neutrofil* dan *eosinofil*, sel darah lain adalah *limfosit* dan *limfoblast* yang mempunyai kedekatan ciri. Jika dilihat berdasarkan kondisi nyata di dalam tubuh manusia, hasil ekstraksi pada beberapa sel tersebut memang cukup mampu merepresentasikan bentuk identik sel. Tetapi pada aplikasi ini, kedekatan ciri antar sel dapat membuat algoritma menjadi tidak akurat. Karena

algoritma ini hanya menggunakan jarak tetangga terdekat dan penentuannya dengan mayoritas tetangga yang paling banyak. Ciri-ciri yang digunakan dalam penelitian ini tidak dapat merepresentasikan bentuk identik sel, sehingga tidak dapat membedakan jenis sel dengan baik dan benar.

Akurasi aplikasi ini dapat dihitung dengan menggunakan *confusion matrix* (tabel kebingungan sistem). Tabel kebingungan sistem merupakan proses dalam melakukan pencocokan hasil klasifikasi pada sistem ini dengan data yang akurat [10]. Data yang digunakan sebagai acuan untuk pencocokan hasil klasifikasi adalah sebagai berikut:

Basofil :0; Neutrofil : 3; Eosinofil :4; Limfosit : 21; Limfoblast : 62; Monosit : 3; Tidak terdeteksi :7

Rekapitulasi hasil pengujian menggunakan *confusion matrix* ditunjukkan pada Tabel 5.

Tabel 5 *Confusion Matrix*

Real/Sistem	Basofil	Neutrofil	Eosinofil	Limfosit	Limfoblast	Monosit	Tidak Terdeteksi
Basofil	0	0	0	0	0	0	0
Neutrofil	0	0	0	3	0	0	0
Eosinofil	0	0	0	4	0	0	0
Limfosit	0	0	0	19	0	2	0
Limfoblast	2	1	3	8	44	4	0
Monosit	0	0	0	2	0	1	0
Tidak terdeteksi	1	1	0	2	3	0	0

Berdasarkan data yang diperoleh pada Tabel 5 maka dapat ditentukan kategori sel darah yang terdeteksi benar dan salah, dimana:

TN : Jumlah citra yang benar diklasifikasi bukan sel darah

TP : Jumlah citra yang benar diklasifikasi sebagai sel darah

FP : Jumlah citra yang salah diklasifikasi sebagai sel darah

FN : Jumlah citra yang salah diklasifikasi sebagai bukan sel darah

Pada tabel tersebut tidak terdapat kategori TN (*True negative*) karena tidak ada citra yang benar diklasifikasi sebagai bukan sel darah. Pada tabel kebingungan sistem akan menghasilkan akurasi tinggi jika pada diagonalnya berisi nilai jumlah sel darah putih yang diklasifikasi dengan benar. Dengan menggunakan tabel kebingungan sistem tersebut, maka dapat dihitung presentase akurasi dari pengujian klasifikasi dengan menggunakan persamaan (1).

$$\% \text{ Akurasi} = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (1)$$

Sehingga perhitungan tersebut menjadi:

$$\begin{aligned} \% \text{ Akurasi} &= \frac{64 + 0}{64 + 29 + 7 + 0} \times 100\% \\ &= 64\% \end{aligned}$$

Jadi presentase akurasi untuk proses klasifikasi sel darah putih adalah sebesar 64%. Tingkat akurasi proses klasifikasi ini dipengaruhi oleh proses *thresholding* pada sel darah putih. Kekurangan dari proses *thresholding* pada sistem ini yaitu nilai *threshold* hanya dapat melakukan *thresholding* pada citra dengan warna gelap, sehingga pada beberapa citra sel darah putih yang mempunyai *sitoplasma* dengan warna terang tidak dapat dilakukan *thresholding* secara penuh yang berarti bahwa ciri luas, *perimeter* dan kebundaran menjadi tidak akurat karena sisi sel darah putih tidak tersegmentasi dengan baik. Dengan presentase akurasi tersebut maka metode *k-nearest neighbor* (knn) dengan menggunakan *hough circle* dan proses *thresholding* kurang baik digunakan untuk klasifikasi sel darah putih.

4. KESIMPULAN

Telah berhasil dilakukan klasifikasi sel darah putih dengan metode *k-nearest neighbor* (knn) dan diperoleh kesimpulan sebagai berikut :

1. Metode segmentasi sel darah putih dari citra darah yang dikembangkan mampu mengidentifikasi sel darah putih mempunyai akurasi sebesar 78%.
2. Metode klasifikasi citra sel darah putih dengan *k-nearest neighbor* (knn) mampu mengklasifikasi jenis sel darah putih secara keseluruhan. Dengan menggunakan *confusion matrix*, metode mempunyai akurasi sebesar 64%.

5. SARAN

Berikut saran-saran yang dapat dilakukan untuk pengembangan dalam penelitian selanjutnya yang sejenis :

1. Metode segmentasi citra perlu ditingkatkan akurasinya dengan mencari nilai parameter-parameter *hough circle* atau dengan metode lain karena mempertimbangkan metode ini yang kurang baik dalam melakukan segmentasi citra.
2. Perlu menambahkan ekstraksi ciri agar dapat merepresentasikan jenis sel darah putih misalnya tidak hanya berdasarkan bentuk dan warna, tetapi bisa ditambahkan dengan ekstraksi ciri berdasarkan tekstur.
3. Menambahkan sampel jenis sel darah yang jumlahnya sedikit misalnya *basofil*, *eosinofil* dan *monosit*.
4. Metode klasifikasi perlu ditingkatkan akurasinya dengan metode lain misalnya jaringan syaraf tiruan.

DAFTAR PUSTAKA

- [1] Hiremath, P.S., Bannigidad, P., Geeta, S. 2010. *Automated Identification and Classification of White Blood Cells (Leukocytes) in Digital Microscopic Images*. IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR, 2010 Halaman 59. Dept. of Computer Science, Gulbarga University, Gulbarga, Karnataka, India.
- [2] Wiyanti, A, 2013, *Multilayer Perceptron Network Clasification Of White Blood Cell's Components With Multilayer Perceptron Network*, Jurnal Digilib ITS, Surabaya.
- [3] Ramirez-Cortes, J.M., Gomez-Gil, P., Alarcon- Aquino, V., Gonzalez-Bernal, J., Garcia-Pedrero, A., 2011, *Neural Networks and SVM-Based Classification of Leukocytes Using the Morphological Pattern Spectrum*, Department of Electronics, National Institute of Astrophysics, Optics and Electronics, Luis Enrique Erro No. 1 Tonantzintla, Puebla, 72840, Mexico.
- [4] Widiarsana, I.G.A., 2011, *Metode Klasifikasi K-Nearest Neighbor (KNN)*, Fakultas Teknik, Universitas Udayana, Denpasar.
- [5] Simon, E, 2014, *Penerapan Algoritma Jaringan Syaraf Tiruan Propagasi Balik dan ransformasi Hough untuk Deteksi Lokasi Mata Pada Citra Digital*, Program Studi Teknik Informatika, STIMIK GI MDP, Palembang.
- [6] Linda, A, 2012, *Penerapan Region of Interest (ROI) pada Metode Kompresi JPEG200*, Departemen Teknik Informarika, Institut Teknologi Bandung, Bandung.
- [7] Ahmad, U, 2005, *Pengolahan Citra Digital*, Yogyakarta: Graha Ilmu
- [8] Putra, D, 2010, *Pengolahan Citra Digital*, Yogyakarta: Penerbit Andi
- [9] Katz, A.R.J. 2000, *Image Analysis and Supervised Learning in the Automated Differentiation of White Blood Cells from Microscopic Images*, Department of Computer Science, RMIT
- [10] Krisandi, N, 2013, *Algoritma K-Nearest Neighbor dalam Klasifikasi Data Hasil Produksi Kelapa Sawit Pada PT. Minamas Kecamatan Parindu*, *Buletin Ilmiah Math. Stat. dan Terapannya (Bimaster)*, No.1, Volume 02, halaman 33-38.