

Penggabungan Fitur Tekstur yang *Invariant* terhadap Iluminasi dan Fitur Bentuk untuk Deteksi *Acute Lymphoblastic Leukemia*

Rizal A. Saputra¹, Chastine Faticah², Nanik Suciati³

Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Institut Teknologi Sepuluh Nopember Surabaya

Kampus ITS Sukolilo Jl. Raya ITS, Surabaya-60111, Jawa Timur

E-mail: ¹rizaladisaputra@gmail.com., ²chastine@if.its.ac.id, ³nanik@if.its.ac.id

Masuk: 6 Juli 2015; Direvisi: 15 Juli 2015; Diterima: 27 Juli 2015

Abstract. Detection with microscopic blood image can help early detection of Acute Lymphoblastic Leukemia (ALL). Therefore, image acquisition process under lighting variation cause varying illumination image, so it's needed to find texture feature extraction method that is invariant towards illumination. Shape feature also needed in this study because can represent characteristics of microscopic blood image. This study proposes combination of texture feature that is illumination invariant and shape feature for ALL detection. Texture feature will be extracted using Complete Robust Local Binary Pattern (CRLBP) method and will be tested on microscopic blood image dataset named ALL_IDB1. Testing will be conducted by using various combination of different texture feature and shape feature. Combination of shape feature and CRLBP is perform better than others. In individual cell test, highest result using SVM Linear with accuracy 90.89%, sensitivity 94.24% and specificity 64.82%. Classification using ALL image reach accuracy 88.00 %, sensitivity 82.35% and specificity 100%.

Keywords: Acute Lymphoblastic Leukemia detection, Complete Robust Local Binary Pattern, Local Binary Pattern, shape feature, texture feature.

Abstrak. Deteksi dengan citra mikroskopik sel darah dapat membantu untuk deteksi dini Acute Lymphoblastic Leukemia (ALL). Namun, proses akuisisi citra mikroskopik dengan variasi pencahayaan yang berbeda menyebabkan iluminasi citra menjadi beragam sehingga dibutuhkan metode yang dapat mengekstraksi fitur tekstur yang invariant terhadap iluminasi. Fitur bentuk juga dibutuhkan dalam penelitian ini karena dapat merepresentasikan perbedaan pada citra mikroskopik sel darah. Penelitian ini mengusulkan penggabungan fitur tekstur yang invariant terhadap iluminasi dan fitur bentuk untuk deteksi dini ALL. Fitur tekstur akan diekstraksi dengan menggunakan metode Complete Robust Local Binary Pattern (CRLBP) dan diuji coba pada dataset ALL_IDB1. Uji coba dilakukan dengan variasi penggabungan fitur bentuk dan fitur tekstur. Penggabungan fitur bentuk dan CRLBP merupakan kombinasi fitur dengan performansi paling baik. Pada pengujian sel tunggal memberikan hasil tertinggi pada klasifikasi SVM Linear dengan akurasi 90,89%, sensitifitas 94,24% dan spesifisitas 64,82%. Pada klasifikasi citra ALL akurasi mencapai 88,00%, dengan sensitifitas 82,35% dan spesifisitas 100%.

Kata Kunci: Complete Robust Local Binary Pattern, deteksi Acute Lymphoblastic Leukemia, Local Binary Pattern, fitur bentuk, fitur tekstur

1. Pendahuluan

Penyakit leukimia merupakan penyakit kanker darah yang disebabkan produksi sel darah putih berlebih dan tidak terkendali sehingga menyebabkan fungsi normal darah menjadi terganggu. *Acute Lymphoblastic Leukemia* (ALL) atau yang biasa juga dikenal dengan *acute lymphocytic leukemia* merupakan merupakan kanker yang paling umum diderita anak-anak. Setiap tahun, sekitar satu dari 50.000 orang terdiagnosis ALL dengan 70% di antaranya anak-anak (Paolini, dkk., 2011). Sebanyak 80% dari anak-anak yang terdiagnosa ALL dapat sembuh

setelah menjalani perawatan dan pengobatan. Namun untuk pasien dewasa persentasi tersebut menurun menjadi 40%-50% (Bassan, dkk., 2009; Bassan dan Hoelzer, 2011). Oleh karena itu deteksi dini penting dilakukan untuk pemulihan pasien khususnya pada anak-anak.

Deteksi ALL secara manual oleh tenaga ahli medis cenderung kurang presisi dan membutuhkan waktu yang lama. Oleh karena itu beberapa penelitian telah menggunakan teknik pengolahan citra untuk deteksi ALL. ALL didiagnosis dengan cara deteksi sel darah putih untuk dipisahkan dari citra keseluruhan dan kemudian dilakukan ekstraksi fitur lalu diklasifikasikan. Untuk melakukan deteksi ALL melalui citra mikroskopik darah fokus fitur yang menjadi pembeda adalah fitur bentuk dan fitur tekstur. Fitur tekstur dapat mendeskripsikan distribusi kromatin pada nukleus sedangkan fitur bentuk dapat merepresentasikan perbedaan morfologi sel darah.

Fokus ekstraksi fitur bentuk salah satunya dengan memanfaatkan fitur morfologi (Scotti, 2005; Raje dan Rangole, 2014). Namun sebagian besar peneliti menggabungkan berbagai fitur karena fitur bentuk saja tidak cukup mewakili fitur sel darah putih. Contoh penggabungan fitur yakni dengan menggabungkan fitur bentuk (dimensi fraktal), tekstur (*Haushdorff Dimension* dan *Contour Signature*) serta fitur warna (Mohapatra dan Patra, 2010), fitur tekstur (*Gray Level Statistic* dan *Matrix Co-Occurance*), bentuk, dan warna (Mohapatra, dkk., 2011), serta fitur bentuk dan fitur tekstur GLCM (Madhloom, 2012). Penelitian berikutnya banyak difokuskan pada ekstraksi fitur tekstur untuk deteksi *Acute Lymphoblastic Leukemia* (ALL) menggunakan Wavelet, GLCM, dan Transformasi Fourier (Mohapatra, dkk., 2014) serta menggunakan GLCM (Putzu, 2014).

Database citra mikroskopik sel darah ALL_IDB1 (Scotti, dkk., 2005) yang digunakan pada penelitian ini diakuisisi dengan menggunakan sensor kamera yang berbeda dan variasi pencahayaan yang berbeda pula saat akuisisi. Hal ini menyebabkan citra mikroskopik sel darah yang ada pada *database* ALL_IDB1 memiliki iluminasi yang berbeda-beda. Secara umum, rata-rata hasil klasifikasi pada *database* dengan iluminasi yang beragam menurun 25% dibandingkan pada *database* yang iluminasinya seragam (Mäenpää dan Pietikäinen, 2004). Hanbury (2005) menyebutkan perubahan iluminasi tidak merubah isi dari tekstur namun mempengaruhi intensitas tekstur dan sebagian besar dari analisis tekstur sensitif terhadap perubahan iluminasi. Sehingga untuk mengatasi masalah iluminasi beragam pada citra akan lebih baik menggunakan ekstraksi fitur tekstur pada citra keabuan. Ojala, dkk. (1996), mengusulkan penggunaan *Local Binary Pattern* (LBP) untuk klasifikasi tekstur. LBP merupakan operator yang efisien untuk mendeskripsikan pola lokal sebuah citra dan mencapai hasil yang optimal untuk klasifikasi tekstur pada beberapa *database* (Ojala, dkk., 2002). LBP konvensional dan turunannya telah banyak digunakan untuk deskriptor fitur termasuk untuk deteksi ALL (Singhal dan Singh, 2014).

Pola biner dari LBP menurut Nanni, dkk. (2010) adalah fitur tekstur dengan komputasi sederhana sehingga cocok sebagai deskriptor tekstur untuk citra medis. Akan tetapi, performa LBP konvensional pada citra dengan iluminasi beragam masih kurang stabil. Guo, dkk. (2010) melakukan pengembangan LBP dengan mengusulkan *Complete LBP* (CLBP) yang menggabungkan LBP dengan *Local Different Sign-Magnitude Transform* (LDSMT). Walaupun CLBP dapat menyelesaikan masalah pola yang berbeda, namun tidak semua pola itu dapat dibedakan. Contohnya adalah dengan nilai piksel atau pola yang berbeda namun memiliki kode CLBP yang sama. Zhao, dkk. (2013) mengusulkan *framework* baru yang dinamakan *Complete Robust LBP* (CRLBP) untuk mengatasi kekurangan tersebut. CRLBP menggunakan *Weighted Local Gray-Level* (WLG) yang merepresentasikan nilai level keabuan dari tekstur lokal untuk mengganti nilai *threshold*. CRLBP juga mengadaptasi penelitian Guo, dkk. sehingga fitur yang dihasilkan ada tiga yaitu RLBP, RLBP_M dan RLBP_C. Hasil dari CRLBP diujicobakan pada data dengan iluminasi beragam menunjukkan metode tersebut dapat menunjukkan performa lebih stabil pada *invariant* iluminasi daripada CLBP.

Pada penelitian yang dilakukan oleh Putzu (2014) untuk klasifikasi ALL dilakukan dengan menggunakan *database* ALL_IDB1. Namun penelitian ini hanya menggunakan 33 citra dengan kondisi iluminasi yang sama untuk data uji. Hal ini dikarenakan ekstraksi fitur sangat

erat hubungannya dengan pencahayaan, karena pencahayaan dapat merubah intensitas tekstur sehingga dapat mempengaruhi hasil deteksi ALL. Oleh karena itu, penelitian ini mengusulkan penggabungan fitur tekstur yang *invariant* terhadap iluminasi dan fitur bentuk untuk deteksi penyakit *Acute Lymphoblastic Leukemia* (ALL). Fitur bentuk yang digunakan terdiri fitur rasio ukuran serta fitur geometri dan fitur tekstur menggunakan *Complete Robust Local Binary Pattern* (CRLBP). Penggabungan fitur ini diharapkan akan menghasilkan fitur yang dapat meningkatkan hasil akurasi untuk deteksi penyakit ALL.

2. Tinjauan Pustaka

2.1. Fitur Bentuk

Secara umum representasi suatu objek melalui fitur bentuk dibagi menjadi dua yakni *boundary descriptor* dan *region descriptor* (Nixon dan Aguado, 2012). Untuk membedakan citra mikroskopik sel darah, fitur bentuk memegang peranan penting karena dapat mendeskripsikan struktur dan ukuran dari leukosit. *Area* dan *perimeter* dari leukosit adalah fitur umum yang digunakan untuk merepresentasikan ukuran dari leukosit. Sedangkan secara struktur dapat menggunakan *region* maupun *boundary descriptor*. Untuk ekstraksi bentuk pada citra mikroskopik sel darah dapat dibagi atas dua yaitu dengan menggunakan rasio ukuran dan geometri (Putzu, dkk., 2014; Faticah, dkk., 2012).

Pada analisis leukosit selain perhitungan fitur geometri dapat ditambahkan dua pengukuran spesifik rasio ukuran dengan memanfaatkan *area* dan *perimeter* (Putzu, dkk., 2014). Ukuran rasio dari *area* antara nukleus dan membran sel dapat didefinisikan dengan permukaan luas area atau sama dengan jumlah piksel dari masing-masing *binary mask* yakni S_{nu} yang merepresentasikan nukleus dan S_{cell} untuk representasi sel. Selain itu untuk *perimeter* antara nukleus dan sel dapat didefinisikan sebagai rata-rata perimeter atau sama dengan jumlah batas piksel (*boundary pixel*) dari masing masing *binary mask* yakni P_{nu} dan P_{cell} (Faticah, dkk., 2012). Rasio ukuran dari *area* dan *perimeter* dirumuskan pada Persamaan 1 dan Persamaan 2, dengan S_{nu} adalah *area* nukleus, S_{cell} adalah *area* sel, P_{nu} adalah *perimeter* nukleus, dan P_{cell} adalah *perimeter* sel. Adapun fitur geometri yang digunakan meliputi *area*, *perimeter*, *solidity*, *eccentricity*, *equivdiameter*, *circularity*, dan *roundness*.

$$NCR_{area} = \frac{S_{nu}}{S_{cell}} \quad (1)$$

$$NCR_{perimeter} = \frac{P_{nu}}{P_{cell}} \quad (2)$$

2.2. Complete Robust Local Binary Pattern (CRLBP)

Ojala, dkk. (1996) mengajukan penggunaan *Local Binary Pattern* (LBP) untuk klasifikasi tekstur. LBP adalah operator simpel namun efisien dalam mendeskripsikan pola lokal pada citra, dan telah berhasil menunjukkan performa yang baik dalam klasifikasi. LBP juga telah diadaptasi pada banyak aplikasi, seperti pengenalan wajah, pengenalan tekstur dinamis, *shape localization* dan juga pada berbagai citra medis. Jika diberikan sebuah piksel dalam sebuah citra, maka rumus LBP dapat dihitung melalui perbandingan ke tetangganya seperti yang terlihat pada Persamaan 3, dengan g_c adalah nilai keabuan pada piksel tengah, g_p adalah nilai piksel tetangganya, P adalah tetangga yang terlibat, dan R adalah nilai radius ke tetangganya (Ojala, dkk., 2002). Setiap piksel tetangga dilabeli nilai satu atau nol. Jika nilai tetangganya lebih besar dari nilai tengahnya maka akan dilabeli nol dan jika sebaliknya akan diberi nilai satu. Sehingga *threshold* ketetanggaan merepresentasikan kode biner yang menggantikan nilai tengah. Kode biner ini akan dikonversi ke angka desimal dengan cara mengalikan nilai biner dengan bobotnya.

$$LBP_{P,R} = \sum_{p=0}^{p-1} s(g_p - g_c) 2^p, s(x) = \begin{cases} 1, x \leq 0 \\ 0, x > 0 \end{cases} \quad (3)$$

Untuk meningkatkan kemampuan diskriminatif struktur lokal dari LBP, Guo, dkk. (2010) mengusulkan metode yang diberi nama *Complete Local Binary Pattern* (CLBP). Pada CLBP citra asli direpresentasikan sebagai *center gray level* (c) dan perbedaan lokal citra didekomposisi ke dalam dua komponen pengganti yakni *sign* (S_p) dan *magnitude* (m_p) dengan menggunakan *Local Difference Sign- Magnitude Transform* (LDSMT). Selanjutnya fitur-fitur tersebut diberi nama CLBP-*Sign* (CLBP_S), CLBP-*Magnitude* (CLBP_M), dan CLBP-*Center* (CLBP_C). Untuk menyeimbangkan informasi pada piksel individual maka Zhao, dkk. (2013) mengusulkan penggunaan *Weighted Local Gray Level* (WLG) untuk mengganti nilai keabuan dari piksel pusat yang ditunjukkan pada Persamaan 4, dengan g merepresentasikan nilai keabuan dari piksel tengah dan $g_i=(i=1,\dots,8)$ menunjukkan nilai keabuan pada piksel tetangga dan α parameter yang ditentukan oleh *user*.

Robust Local Binary Pattern (RLBP) mengganti nilai piksel pusat dari LBP dengan nilai hasil perhitungan WLG. Sehingga didapatkan Persamaan 5 (Zhao, dkk., 2013), dengan g_c , g_p , P dan R seperti didefinisikan pada Persamaan 3 dan $g_i=(i=1,\dots,8)$ menunjukkan nilai keabuan dari piksel ketetangga g_c . Untuk menyelesaikan pola yang membingungkan pada LBP, RLBP mengadaptasi kerangka kerja dari CLBP. Sehingga nilai m_p dapat dirumuskan seperti pada Persamaan 6 (Zhao, dkk., 2013), dengan g_c , g_p , P dan R seperti didefinisikan pada Persamaan 3, dimana $g_i=(i=1,\dots,8)$ menunjukkan nilai keabuan dari piksel ketetangga g_c dan α nilai parameter pada WLG. RLBP-*Magnitude* mengukur nilai lokal varian dari WLG, sehingga dapat diformulasikan seperti pada Persamaan 7 (Zhao, dkk., 2013), dengan *threshold* c menunjukkan nilai rata-rata m_p dari keseluruhan citra.

Titik pusat citra yang mewakili nilai keabuan titik pusat, juga memiliki informasi yang signifikan sehingga Zhao, dkk. (2013) membuat operator RLBP-*Center* (RLBP_C) untuk mengekstraksi informasi lokal yang berada di titik pusat. Adapapun RLBP_C dapat dirumuskan seperti pada Persamaan 8, dengan *threshold* c_l menunjukkan nilai *average local gray-level* dari keseluruhan citra. Mengadaptasi dari penelitian Guo, dkk. (2010), Zhao, dkk. (2013) mengkombinasikan ketiga fitur RLBP, RLBP_M, dan RLBP_C sebagai CRLBP.

$$WLG = \frac{\sum_{i=1}^8 g_i + \alpha g}{8 + \alpha} \quad (4)$$

$$RLBP_{P,R} = \sum_{p=0}^P s(g_p - WLG_c) 2^p \quad (5)$$

$$m_p = \left| \frac{\sum_{i=1}^8 g_{pi} + \alpha g_p}{8 + \alpha} - \frac{\sum_{i=1}^8 g_{ci} + \alpha g_c}{8 + \alpha} \right| \quad (6)$$

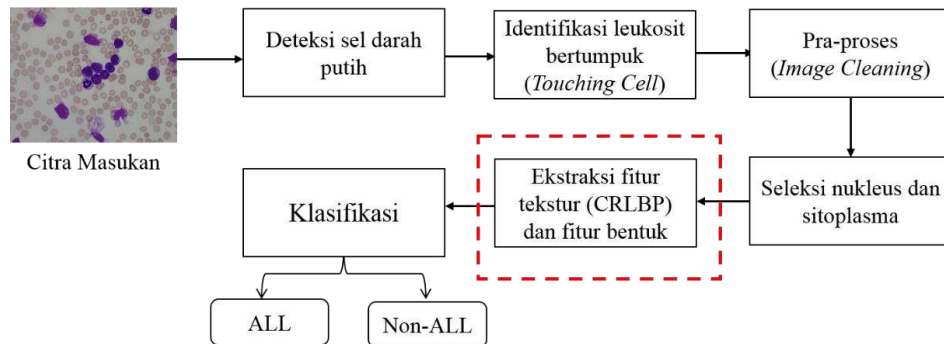
$$RLBP_M_{P,R} = \sum_{p=0}^{P-1} s(m_p - c) 2^p \quad (7)$$

$$RLBP_C_{P,R} = s(WLG_c - c_l) \quad (8)$$

3. Metodologi Penelitian

Secara garis besar, perancangan sistem deteksi yang diajukan terdiri atas enam proses, yaitu pengumpulan *dataset*, deteksi sel darah putih, pra-proses, seleksi nukleus dan sitoplasma, ekstraksi fitur, dan klasifikasi seperti yang ditunjukkan pada Gambar 1. *Dataset* yang digunakan merupakan citra mikroskopik sel darah yang memiliki beragam iluminasi. Komponen sel darah yang digunakan adalah sel darah putih sehingga perlu dilakukan deteksi sel darah putih dengan menggunakan Algoritma Zack. Selanjutnya dilakukan identifikasi sel bertumpuk dengan menggunakan analisis *roundness* sehingga didapatkan sel limfosit tunggal. Komponen sel darah putih lainnya akan dibersihkan pada tahap pra-proses dengan menggunakan analisis *solidity* yang hasilnya kemudian akan dikorespondensikan ke citra RGB untuk di-*cropping*. Sel-sel tunggal tersebut kemudian dilakukan seleksi nukleus dan sitoplasma. Hasil citra biner kemudian diekstraksi fitur dengan menggunakan fitur bentuk dan citra keabuan diekstraksi fitur tekstur

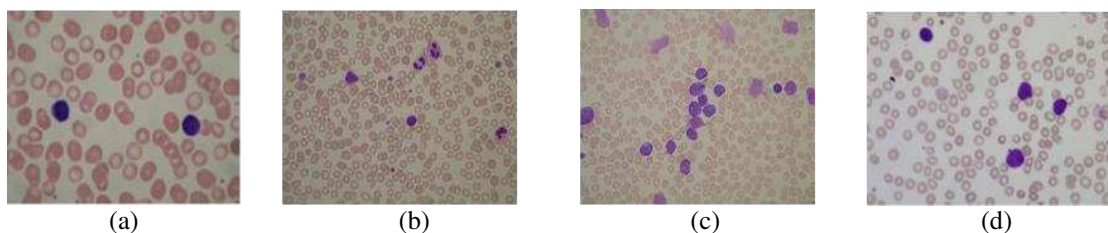
menggunakan *Complete Robust Local Binary Pattern* (CRLBP) yang kemudian nilai fiturnya digabungkan dan diklasifikasi menggunakan *Support Vector Machine* (SVM).



Gambar 1. Rancangan Sistem Klasifikasi Acute Lymphoblastic Leukemia

3.1. Dataset

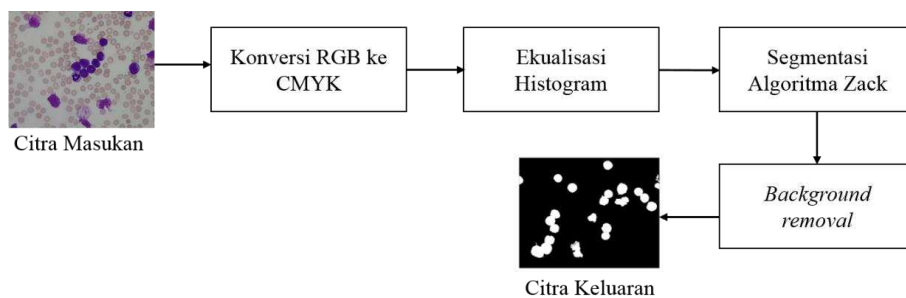
Dataset yang digunakan merupakan dataset ALL_IDB1 yang diakuisisi dan dikumpulkan oleh ahli dari M.Tettamanti Research Centre for Childhood Leukemia and Haemato-Logical Diseases, Monza, Italia dan selanjutnya dapat diunduh pada database Department of Information Technology-Università degli Studi di Milano (Scotti, dkk., 2005). Dataset ini dapat diunduh di <http://www.dti.unimi.it/fscotti/all/> dengan memohon persetujuan terlebih dahulu kepada pihak terkait. Data dari ALL-IDB1 yang digunakan memiliki beragam iluminasi dikarenakan tidak meratanya pencahayaan saat akuisi dan juga panjang sensor dari kamera yang digunakan berbeda sehingga mempengaruhi intensitas cahaya yang diterima. Perbedaan iluminasi pada citra input ditunjukkan pada Gambar 2.



Gambar 2. Dataset ALL_IDB1 dalam variasi iluminasi (a) (b) citra darah individu sehat (c) (d) citra darah individu terdeteksi ALL (Scotti, dkk., 2005)

3.2. Deteksi Sel Darah Putih

Pendekatan yang digunakan untuk deteksi sel darah putih pada penelitian ini mengadopsi pendekatan yang diusulkan oleh Putzu, dkk. (2014). Berbeda dengan pendekatan lain yang melakukan segmentasi sel darah putih dengan mendeteksi nukleus pertama kali, pendekatan yang digunakan oleh Putzu melakukan segmentasi sel darah putih dengan terlebih dahulu mendeteksi membran sel atau keseluruhan sel darah putih. Secara umum proses untuk deteksi sel darah putih dapat dilihat pada Gambar 3.



Gambar 3. Rancangan Sistem Deteksi Sel Darah Putih

3.3. Identifikasi Limfosit Tunggal

Untuk mengidentifikasi sel individual dan sel yang bersentuhan dilakukan analisis *roundness*. Analisis *roundness* merupakan pengukuran seberapa bulat bentuk sel. Sel individual akan memiliki bentuk lebih bulat (*round*) dibandingkan sel yang bersentuhan. *Roundness* bernilai satu untuk objek yang bundar dan kurang dari satu untuk objek yang menjauhi bentuk bundar; pengukuran ini relatif insensitif terhadap *boundary* yang *irregular*. Komponen yang terhubung (*connected component* atau *region*) dengan *roundness* lebih besar dari nilai *threshold* diklasifikasikan sebagai sel darah putih individual, selanjutnya diproses ke tahap analisis berikutnya. Sedangkan *region* yang bernilai *roundness* kurang dari nilai *threshold*, digolongkan sebagai sel yang bersentuhan atau grup sel dan akan diproses pada tahapan selanjutnya

3.4. Pra-Proses

Pra-proses pada penelitian ini fokus pada *image cleaning* yakni penghapusan semua sel darah putih yang terdapat pada tepi citra dan semua komponen abnormal (non-sel darah putih). Tahapan ini bertujuan untuk mengurangi *error* pada tahap analisis akhir. Pertama, ukuran *area* dan *convex area* dihitung untuk masing-masing sel darah putih. *Area* dan *convex area* akan digunakan untuk menghitung *solidity* atau kepadatan objek yang akan digunakan untuk memisahkan komponen abnormal. Selanjutnya akan dihitung nilai *threshold* dari citra. Komponen yang nilai *solidity*-nya di bawah nilai *threshold* akan dianggap sebagai komponen abnormal dan dibuang. Setelah dilakukan proses *image cleaning* citra hasil pra-proses akan dikorespondesikan dengan citra RGB aslinya untuk dilakukan identifikasi sel individual.

3.5. Seleksi Nukleus dan Sitoplasma

Tahapan ini diawali dengan melakukan seleksi sel darah putih tunggal menggunakan *bounding box*, yaitu *rectangle* terkecil yang mencakup komponen terhubung, bertujuan untuk mengisolasi sel darah putih tunggal atau individual sehingga menjadi sub-sub citra. Citra biner nukleus diperoleh dari komponen *S* pada ruang warna HSV melalui operasi *threshold*. Selanjutnya dilakukan operasi pembersihan tepi citra dengan membersihkan citra tepi yang memiliki *area region* kurang dari rata-rata sel darah putih, objek yang bukan ROI pun ikut dihilangkan. Selanjutnya, setelah citra biner nukleus terekstrak, sitoplasma dapat diperoleh dengan pengurangan antara citra biner membran sel dengan citra nukleus. Operasi morfologi dilasi dan erosi digunakan untuk menghubungkan komponen terhubung yang masih terputus.

3.6. Ekstraksi Fitur

Berdasarkan citra hasil pra-proses akan dilakukan perhitungan fitur bentuk. Fitur bentuk akan diekstraksi dari citra biner sedangkan fitur tekstur dengan menggunakan citra level keabuan. Fitur bentuk yang dihitung meliputi dua fitur rasio ukuran dan 14 fitur geometri. Fitur rasio ukuran yang dihitung adalah *area* dan *perimeter* antara nukleus dan keseluruhan sel. Delapan fitur geometri yang dihitung yaitu *area*, *perimeter*, *solidity*, *eccentricity*, *equivdiameter*, *circularity*, dan *roundness* yang masing-masing merepresentasikan geometri dari nukleus dan sitoplasma sehingga akan ada 14 fitur geometri. Untuk menghitung fitur tekstur dilakukan dengan menggunakan metode *Complete Robust Local Binary Pattern* (CRLBP) yang akan menghasilkan tiga nilai fitur yakni RLBP, RLBP_M, dan RLBP_C. Selanjutnya hasil dari ekstraksi fitur bentuk dan tekstur ini akan digabungkan sebagai fitur untuk deteksi penyakit *Acute Lymphoblastic Leukemia* (ALL). Total keseluruhan akan ada 19 jenis fitur yang digunakan.

3.7. Klasifikasi *Acute Lymphoblastic Leukemia*

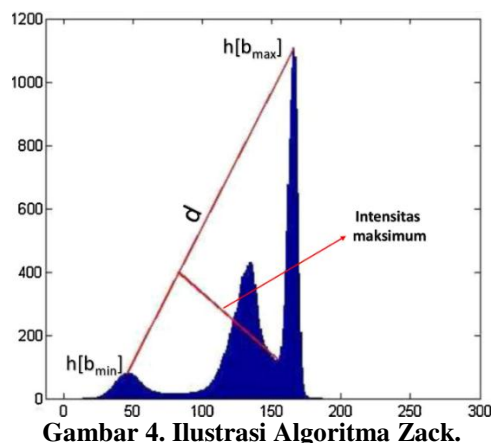
Deteksi penyakit *Acute Lymphoblastic Leukemia* (ALL) diuji dengan cara klasifikasi menggunakan *Support Vector Machine* (SVM). Pada dasarnya SVM merupakan jenis pengklasifikasi untuk dua kelas karena metode ini mengoptimalkan *margin* antar kelas. Uji coba klasifikasi secara garis besar akan dibagi menjadi dua tahap yaitu pengujian klasifikasi per-sel

dan yang kedua akan dilakukan pengujian klasifikasi per citra. Untuk klasifikasi pada sel tunggal termasuk ALL atau Non-ALL akan dilihat dari hasil sel tunggal tersebut saja. Sedangkan untuk klasifikasi pengujian citra, seluruh sel tunggal dalam citra akan diklasifikasi. Citra akan terklasifikasi sebagai ALL jika paling tidak ada satu sel tunggal yang terklasifikasi sebagai ALL (Labati, dkk., 2011).

Adapun skenario percobaan pada masing-masing pengujian sel dan citra akan dilakukan kombinasi ekstraksi fitur yang meliputi: (1) fitur bentuk, (2) fitur tekstur LBP (3) gabungan fitur bentuk dan LBP, (4) fitur tekstur CRLBP, dan (5) gabungan fitur bentuk dan CRLBP. Pada pengujian klasifikasi sel, metode klasifikasi yang digunakan adalah *Support Vector Machine* dengan tiga kernel yakni Linear, *Polynomial 2* dan *Radial Basis Function* (RBF). Sedangkan pada pengujian klasifikasi citra hanya akan menggunakan SVM dengan salah satu kernel yang memberikan hasil performansi paling tinggi berdasarkan pengujian sel. *Cross validation* digunakan untuk validasi keakuratan dalam pengujian penelitian ini dalam penelitian ini dikarenakan data yang tidakimbang pada jumlah sel.

4. Hasil dan Pembahasan

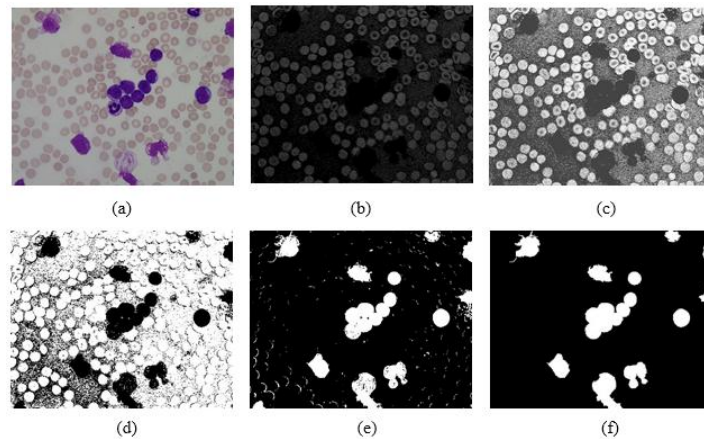
Untuk keperluan penelitian ini, *dataset* yang digunakan berjumlah 98 data citra darah yang terdiri dari 49 citra dengan label terinfeksi ALL dan 49 citra lainnya tidak terinfeksi ALL (Non-ALL). Sebanyak 65 citra merupakan citra yang diakuisisi menggunakan kamera Canon PowerShot G5 RGB 24-bit dengan resolusi 2592x1944 piksel dengan format JPG. Sedangkan 33 citra lainnya diakuisi dengan kamera Olympus C2500L menghasilkan resolusi 1712x1368 piksel. Masing masing citra akan dilakukan deteksi membran sel darah putih dilakukan pada ruang warna CMYK. Membran sel darah putih lebih kontras pada komponen Y dari model CMYK karena warna kuning terdapat pada semua elemen citra, kecuali sel darah putih. Redistribusi *grey level* citra perlu dilakukan agar proses segmentasi lebih mudah sehingga selanjutnya ekualisasi histogram atau *contrast stretching* dapat digunakan. Pada penelitian ini dilakukan segmentasi dengan penentuan nilai *threshold* berdasarkan metode segitiga (*triangle*) atau Algoritma Zack. Algoritma ini digunakan secara efektif karena histogram menunjukkan puncak yang jelas antara puncak yang tinggi dan rendah, seperti yang terdapat pada histogram komponen Y dari sel darah putih dan sel darah merah. Ilustrasi dari Algoritma Zack ditunjukkan pada Gambar 4.



Gambar 4. Ilustrasi Algoritma Zack.

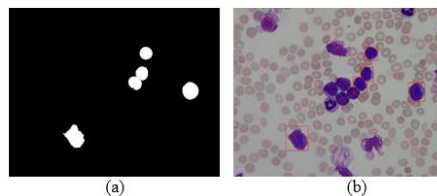
Untuk meningkatkan hasil segmentasi, *background* citra harus dihilangkan. Pada penelitian ini, pendekatan yang digunakan untuk ekstraksi *background* melibatkan penggunaan algoritma *threshold* otomatis berdasarkan komponen *green* atau *G* dari ruang warna RGB. Selanjutnya komponen *green* akan disegmentasi dengan Algoritma Zack dan dilakukan substraksi dengan komponen *Y* dari hasil konversi CMYK yang sudah disegmentasi sebelumnya. Penghilangan komponen selain sel darah putih (*background*) dengan perhitungan *opening area* yakni dengan menghapus objek yang memiliki ukuran lebih kecil dari area

structuring element. *Structuring element* yang digunakan berbentuk bulat dan ukurannya dihitung berdasarkan rata-rata objek (sel darah putih) yang berada pada citra. Hasil proses identifikasi pada sel darah putih dapat dilihat pada Gambar 5.



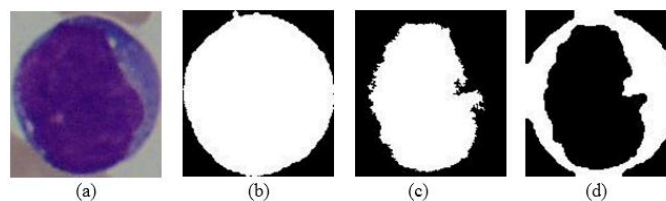
Gambar 5. Identifikasi Sel Darah Putih: (a) citra asli, (b) komponen Y citra, (c) histogram ekualisasi Y, (d) segmentasi Zack Y, (e) hasil opening, (f) hasil final penghilangan *background*.

Setelah penghilangan *background* dilakukan identifikasi limfosit bersentuhan dengan menggunakan analisis *roundness* dengan nilai ambang 0,8. Komponen dengan nilai *roundness* lebih besar dari 0,8 dianggap sebagai limfosit tunggal, sedangkan yang kurang dari 0,8 dianggap sebagai *touching cell*. Selanjutnya akan dilakukan pra-proses yakni pembersihan tepi citra dan komponen lain yang bukan termasuk limfosit dengan menggunakan perhitungan *solidity* dengan nilai ambang 0,87. Komponen dengan nilai ambang dibawah 0,75 akan dihilangkan. Selanjutnya citra akan dikorespondensikan ke citra RGB seperti terlihat pada Gambar 6, untuk dilakukan *cropping* sehingga menjadi beberapa citra sel.



Gambar 6. (a) Citra hasil pra-proses (b) Citra dikorespondensikan ke dalam image RGB

Setelah limfosit tunggal teridentifikasi langkah berikutnya adalah segmentasi untuk mendapatkan nukleus dan sitoplasma dari sel. Pendekatan seleksi nukleus memanfaatkan komponen warna *green* (G) dari citra warna RGB yang kemudian dilakukan *thresholding* dengan metode Otsu. Hasil citra biner dari komponen *green* (G) dikombinasikan dengan hasil biner komponen S dari HSV yang juga telah di-*thresholding* dengan Otsu. Untuk mendapatkan sitoplasma maka citra biner dari citra sel asli akan dikurangkan dengan citra biner nukleus. Selanjutnya akan dilakukan proses morfologi dilasi dan erosi.



Gambar 7. Hasil seleksi nukleus dan sitoplasma (a) Sel individual limfosit, (b) Citra biner ROI sel limfosit, (c) citra biner nukleus, (d) citra biner sitoplasma

Sel biner nukleus dan sitoplasma seperti yang terlihat pada Gambar 7, selanjutnya akan digunakan untuk ekstraksi bentuk sedangkan citra *gray level* dari nukleus akan digunakan untuk ekstraksi fitur. Fitur bentuk yang diekstraksi meliputi dua fitur rasio ukuran dan 14 fitur geometri. Masing-masing tujuh fitur geometri nukleus dan tujuh fitur geometri sitoplasma, sehingga totalnya ada 16 fitur bentuk. Hasil fitur bentuk berupa data matriks berukuran 1x16. Fitur tekstur yang digunakan pada penelitian adalah *Complete Robust Local Binary Pattern* (CRLBP). CRLBP terdiri dari RLBP_Sign (RLBP_S), RLBP_Magnitude (RLBP_M) dan RLBP_Center (RLBP_C). RLBP_S dan RLBP_M masing-masing menghasilkan 256 nilai fitur yang merupakan nilai histogram dari masing-masing fitur tersebut. Sedangkan fitur RLBP_C hanya menghasilkan dua nilai fitur. Sehingga total keseluruhan fitur tekstur dengan menggunakan CRLBP untuk satu sel sebanyak 514. Data fitur berupa matriks berukuran 1x514.

4.1. Pengujian Klasifikasi Sel

Uji coba dilakukan dengan beberapa skenario pengujian. Pengujian pertama yaitu pengujian klasifikasi masing-masing sel limfosit yang terdeteksi (456 sel) dalam tujuh skenario kombinasi fitur dengan menggunakan tiga kernel SVM. Pada klasifikasi sel, data input berupa matriks dari citra yang sudah di ekstraksi berukuran 456 x panjang fitur. Pengujian selanjutnya adalah klasifikasi citra (98 citra) dengan menggunakan klasifikasi SVM dengan kernel terbaik berdasarkan dari pengujian pada tingkat sel. Uji coba dilakukan menggunakan *10-fold* setiap percobaan dan selanjutnya dilakukan penghitungan *confussion matrix*. Input untuk pengujian ini berupa citra RGB.

Tabel 1. Hasil Pengujian Klasifikasi Sel

Kombinasi Fitur	Kernel SVM	Sen (%)	Spe (%)	Acc (%)
Bentuk	Linear	98,14	43,20	82,67
	Poly 2	97,06	43,10	82,89
	RBF	91,96	51,37	86,22
LBP	Linear	91,43	32,20	80,44
	Poly 2	89,19	26,04	81,78
	RBF	87,45	40,00	87,56
Bentuk + LBP	Linear	93,23	60,90	88,44
	Poly 2	91,74	55,24	87,33
	RBF	87,26	30,00	87,33
CRLBP	Linear	93,42	48,55	85,78
	Poly 2	95,56	47,12	85,11
	RBF	86,86	10,00	86,89
Bentuk + CRLBP	Linear	94,24	64,82	90,89
	Poly 2	95,79	54,86	87,11
	RBF	87,26	30,00	87,33

Berdasarkan Tabel 1, hasil dari klasifikasi sel menunjukkan kombinasi fitur bentuk dan CRLBP memiliki performansi paling baik pada SVM dengan kernel linear dengan akurasi 90,89% dan nilai sensitifitas 94,24% serta spesifisitas 64,82%. Walaupun menunjukkan nilai sensitifitas yang tinggi untuk semua skenario pengujian yakni rata-rata diatas 90%, namun spesifisitas pengujian sel ini cukup rendah. Tercatat hanya kombinasi fitur bentuk dan CRLBP pada SVM Linear dengan 64,82% dan kombinasi fitur bentuk dan LBP juga pada SVM Linear dengan 60,90%. Hal ini dikarenakan banyaknya sel ALL yang terklasifikasi sebagai Non-ALL dalam setiap kfold pengujian. Sedangkan jumlah sel Non-ALL yang terklasifikasi benar sebagai Non-ALL juga jumlahnya sedikit dikarenakan adanya ketidakseimbangan data.

Pengujian sel menggunakan klasifikasi SVM dengan menggunakan tiga kernel berbeda yakni Linear, Polynomial 2, dan juga RBF. Tujuannya adalah untuk melihat kernel apakah yang memiliki performansi paling baik yang selanjutnya akan digunakan untuk klasifikasi gambar. Berdasarkan Tabel 1 dapat terlihat bahwa dengan klasifikasi SVM kernel linear menghasilkan akurasi, sensitifitas dan spesifisitas dibandingkan kedua kernel lainnya dengan persentase berturut-turut 90,89%, 94,24% dan 64,82%. Hal ini disebabkan secara default kernel linear di desain untuk klasifikasi dengan dua kelas. Kernel RBF adalah yang paling rendah

performansinya disebabkan kernel RBF lebih didesain untuk *one-class learning* atau *one-class classification* yakni untuk mengidentifikasi objek dari suatu kelas tertentu dari keseluruhan objek dengan *learning* dari data *training* yang berisi objek dari kelas tersebut saja.

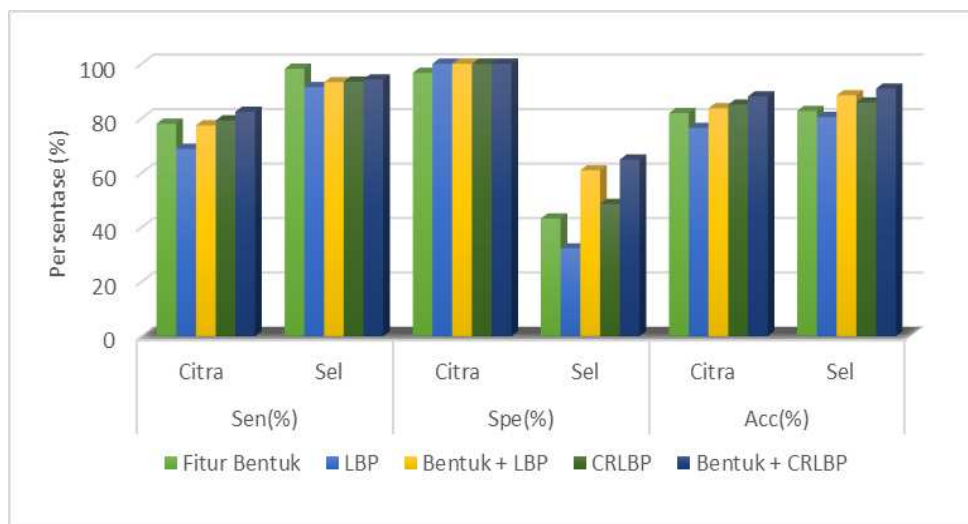
4.2. Pengujian Klasifikasi Citra ALL

Pada pengujian ini setiap citra akan diklasifikasikan satu persatu. Sebanyak 98 citra dengan iluminasi beragam akan menjadi data input dengan menggunakan *kfold* 10. Sehingga pada setiap *fold*-nya akan terdapat 88 atau 89 citra *training* dengan 10 atau sembilan citra uji. Citra akan diklasifikasikan menggunakan *Support Vector Machine* (SVM) kernel linear. Dalam pengujian citra, klasifikasi termasuk ke dalam kelas ALL atau Non-ALL berbeda dibanding dengan pengujian sel. Citra diuji coba pada lima skenario kombinasi ekstraksi fitur yang hasilnya dapat dilihat pada Tabel 2.

Tabel 2. Hasil pengujian klasifikasi citra ALL

Pengujian	Klasifikasi Citra		
	Sen(%)	Spe(%)	Acc(%)
Fitur Bentuk	77,98	96,67	81,89
LBP	68,77	100	76,44
Bentuk + LBP	77,38	100	83,67
CRLBP	79,17	100	85
Bentuk + CRLBP	82,35	100	88

Dari pengujian citra ALL didapatkan nilai akurasi paling tinggi adalah dengan menggunakan kombinasi fitur bentuk dan CRLBP yakni 88,00% dengan sensitifitas 82,35% dan spesifisitas 100%, sedangkan yang terendah adalah dengan menggunakan fitur tekstur LBP saja yakni 76,44% dengan sensitifitas 68,77% dan spesifisitas 100%. Hasil CRLBP lebih tinggi dikarenakan algoritma ini tidak hanya menghitung nilai selisih titik pusat dengan ketetanggan namun juga menghitung nilai *magnitude* dan *center* dari citra serta menggunakan bobot ketetanggan untuk menggantikan nilai asli piksel pusat.



Gambar 8. Perbandingan Hasil Klasifikasi Sel dan Klasifikasi Citra ALL

Nilai akurasi dan sentifitas pada klasifikasi citra lebih rendah dibandingkan pada pengujian klasifikasi sel. Sebaliknya nilai spesifisitas pada klasifikasi citra lebih tinggi dibandingkan klasifikasi sel (Gambar 8). Hal ini dikarenakan perbedaan aturan pada kedua pengujian tersebut. Pada pengujian sel, masing-masing sel akan diklasifikasikan sesuai kelasnya. Jika sel tersebut mengandung blast maka akan teridentifikasi sebagai sel ALL sedangkan jika tidak mengandung blast dikategorikan sebagai Non-ALL, Sedangkan pada klasifikasi citra, jika dalam suatu citra terdapat paling sedikit satu sel saja maka citra tersebut akan dikategorikan sebagai ALL.

4.3. Perbandingan dengan penelitian sebelumnya

Deteksi *Acute Lymphoblastic Leukemia* dengan menggunakan *dataset* ALL-IDB sebelumnya dilakukan oleh Putzu, dkk. (2014) dengan melakukan klasifikasi citra serta Singhal dan Singh (2014) dengan melakukan klasifikasi sel. Metode usulan memiliki hasil yang lebih baik pada klasifikasi sel jika dibandingkan dengan penelitian sebelumnya yang dilakukan oleh Singhal dan Singh. Metode usulan mendapatkan akurasi 90,89%, sedangkan metode klasifikasi sel penelitian sebelumnya hanya mencapai 89,72%. Sedangkan dalam klasifikasi citra metode usulan mencapai akurasi 88,00% dan penelitian sebelumnya mencapai 93,20%. Namun secara analisis hasil penelitian tidak dapat dibandingkan dengan penelitian sebelumnya karena penelitian sebelumnya tidak diuji cobakan pada citra dengan iluminasi beragam sebagaimana dilakukan pada penelitian ini.

5. Kesimpulan

Berdasarkan pengujian, penggabungan fitur bentuk dan fitur tekstur CRLBP dapat digunakan dan menunjukkan performa yang baik untuk klasifikasi *Accute Lymphoblastic Leukemia* (ALL) pada kondisi *database* dengan iluminasi yang beragam. Pada klasifikasi sel didapatkan akurasi 90,89% dengan sensitifity 94,24% dan spesifisitas 64,82% dan pada klasifikasi citra ALL didapatkan akurasi 88,00% dengan sensitifitas 82,35% dan spesifisitas 100%.

6. Future Work

Penelitian berikutnya perlu dilakukan pemisahan sel yang bersentuhan serta pencarian nilai parameter α yang optimal dalam penghitungan nilai level keabuan pada *Complete Robust Local Binary Pattern* (CRLBP) untuk kasus klasifikasi pada citra iluminasi beragam.

Referensi

- Bassan, R., & Hoelzer, D. 2011. Modern Therapy of Acute Lymphoblastic Leukemia. *Journal of clinical oncology*, 29(5), 532-543.
- Bassan, R., Spinelli, O., Oldani, E., Intermesoli, T., Tosi, M., Peruta, B. & Rambaldi, A. 2009. Improved Risk Classification for Risk-Specific Therapy Based on The Molecular Study of Minimal Residual Disease (Mrd) in Adult Acute Lymphoblastic Leukemia (ALL). *BLOOD*, 113(18).
- Fatichah, C., Tangel, M. L., Widyanto, M. R., Dong, F., & Hirota, K. 2012. Parameter Optimization of Local Fuzzy Patterns Based on Fuzzy Contrast Measure for White Blood Cell Texture Feature Extraction. *JACIII*, 16(3), 412-419.
- Guo, Z., & Zhang, D. 2010. A Completed Modeling of Local Binary Pattern Operator for Texture Classification. *Image Processing, IEEE Transactions on*, 19(6), 1657-1663.
- Hanbury, A., Kandaswamy, U., & Adjero, D. A. 2005. Illumination-invariant Morphological Texture Classification. In *Mathematical Morphology: 40 Years On* (pp. 377-386). Springer Netherlands.
- Labati, R. D., Piuri, V., & Scotti, F. 2011. All-IDB: The Acute Lymphoblastic Leukemia Image Database for Image Processing. In *Image processing (ICIP), 2011 18th IEEE international conference on* (pp. 2045-2048). IEEE
- Madhloom, H. T., Kareem, S. A., & Ariffin, H. 2012. A Robust Feature Extraction and Selection Method for the Recognition of Lymphocytes versus Acute Lymphoblastic Leukemia. In *Advanced Computer Science Applications and Technologies (ACSAT), 2012 International Conference on* (pp. 330-335). IEEE.
- Mäenpää, T., & Pietikäinen, M. 2004. Classification with Color and Texture: Jointly or Separately?. *Pattern recognition*, 37(8), 1629-1640.
- Mohapatra, S., & Patra, D. 2010. Automated Leukemia Detection using Hausdorff Dimension in Blood Microscopic Images. In *Emerging Trends in Robotics and Communication Technologies (INTERACT), 2010 International Conference on* (pp. 64-68). IEEE.

- Mohapatra, S., Patra, D., & Satpathy, S. 2014. An Ensemble Classifier System for Early Diagnosis of Acute Lymphoblastic Leukemia in Blood Microscopic Images. *Neural Computing and Applications*, 24(7-8), 1887-1904.
- Mohapatra, S., Samanta, S. S., Patra, D., & Satpathi, S. 2011. Fuzzy Based Blood Image Segmentation for Automated Leukemia Detection. In *Devices and Communications (ICDeCom), 2011 International Conference on* (pp. 1-5). IEEE.
- Nanni, L., Lumini, A., & Brahnam, S. 2010. Local Binary Patterns Variants as Texture Descriptors for Medical Image Analysis. *Artificial intelligence in medicine*, 49(2), 117-125
- Nixon, M., & Aguado, A. S. 2012. *Feature Extraction and Image Processing for Computer Vision*.
- Ojala, T., Pietikäinen, M., & Harwood, D. 1996. A Comparative Study of Texture Measures with Classification Based on Featured Distributions. *Pattern recognition*, 29(1), 51-59.
- Ojala, T., Pietikainen, M., & Maenpaa, T. 2002. Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7), 971-987.
- Paolini, S., Gazzola, A., Sabattini, E., Bacci, F., Pileri, S., & Piccaluga, P. P. 2011. Pathobiology of Acute Lymphoblastic Leukemia. In *Seminars in diagnostic pathology* (Vol. 28, No. 2, pp. 124-134). WB Saunders.
- Putzu, L., Caocci, G., & Di Ruberto, C. 2014. Leucocyte Classification for Leukaemia Detection using Image Processing Techniques. *Artificial intelligence in medicine*, 62(3), 179-191.
- Raje, C., & Rangole, J. 2014. Detection of Leukemia in Microscopic Images Using Image Processing. In *Communications and Signal Processing (ICCSP), 2014 International Conference on* (pp. 255-259). IEEE.
- Scotti, F. 2005. Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Images. In *Computational Intelligence for Measurement Systems and Applications, 2005. CIMSAA. 2005 IEEE International Conference on* (pp. 96-101). IEEE.
- Scotti, F., Labati, R.D., Piuri, V. 2005. *Acute Lymphoblastic Leukemia Image Database for Image Processing*. Department of Information Technology-Università degli Studi di Milano, (Online), (<http://www.dti.unimi.it/fscotti/all/>)
- Singhal, V., & Singh, P. 2014. Local Binary Pattern for Automatic Detection of Acute Lymphoblastic Leukemia. In *Communications (NCC), 2014 Twentieth National Conference on* (pp. 1-5). IEEE.
- Zhao, Y., Jia, W., Hu, R. X., & Min, H. 2013. Completed Robust Local Binary Pattern for Texture Classification. *Neurocomputing*, 106, 68-76.