

Identifikasi Acute Lymphoblastic Leukemia pada Citra Mikroskopis Menggunakan Algoritma Naïve Bayes

Shofiya Shidada^{1*}, Bagus Hariyanto¹

¹ Program Studi Teknik Informatika Fakultas Teknik Universitas Wijaya Kusuma Surabaya
Jl. Dukuh Kupang XXV No.54 Kec. Dukuhpakis Kota Surabaya 60225 Jawa Timur

**shofiya@uwks.ac.id

DOI: <https://doi.org/10.21107/rekayasa.v14i1.9110>

ABSTRACT

Leukemia is a type of blood cancer that occurs when the body overproduces abnormal white blood cells. Acute Lymphoblastic Leukemia (ALL) is a type of acute leukemia. ALL occurs when the spinal cord is excessively producing young lymphocytes, known as lymphoblasts. Leukemia is difficult to detect because it has the same symptoms as other diseases. One way to detect leukemia is to use a complete blood count test. Blood count test is done by calculating the population of red blood cells, white blood cells and platelets. The health condition of the body is indicated by the number of each blood cell. The small number of erythrocyte and abnormal cell shape is indicative of leukemia. How to identify leukemia still using a microscope. In this study the researchers made a way of identifying acute lymphoblastic leukemia cells by image processing, include cropping, segmentation, feature extraction and identification. The method used in image identification is Naïve Bayes Classifier (NBC). NBC is a classification method which applies simple probability calculations using Bayes theorem. The white blood cell image tested using this application will be evaluated with accuracies. The greatest accuracy results from several test scenarios obtained 80% accuracy.

Keyword: acute lymphoblastic leukemia, microscopic image, naïve bayes classifier, thresholding

PENDAHULUAN

Plasma darah dan sel darah merupakan komponen utama pada darah. Jenis sel darah dibedakan menjadi tiga kelompok, yang masing-masing mempunyai tugas tertentu. Sel darah merah disebut eritrosit bertanggung jawab mengedarkan oksigen ke seluruh tubuh. Sel darah putih atau leukosit berfungsi sebagai alarm dalam fungsi kekebalan tubuh. Keping darah atau trombosit berperan dalam proses pembekuan darah. Sedangkan ilmu yang digunakan untuk mempelajari anatomi, fisiologi dan patologi sel darah dikenal dengan hematologi (Noercholis & Muslim, 2013).

Indikasi kondisi kesehatan manusia ditunjukkan oleh kondisi leukosit. Sistem kekebalan tubuh manusia digawangi oleh leukosit. Sel ini bertugas untuk membasmikan kuman atau bahan penyakit. Kondisi leukosit juga merupakan indikator adanya infeksi yang diderita oleh tubuh. Sel darah putih mampu keluar dari pembuluh darah dan berpatroli

untuk menemukan bibit penyakit berupa virus, jamur, bakteri, dan benda asing lainnya. Jumlah leukosit akan meningkat bila tubuh terserang penyakit, selanjutnya kembali normal saat penyakit telah teratas. Tetapi ada sebuah kondisi dimana keberadaan leukosit menjadi ancaman bagi tubuh. Hal ini terjadi bila ada gangguan atau kelaian pada darah. Salah satu gangguan pada darah ini diantaranya adalah pada saat jumlah leukosit berlebihan, sehingga memakan sel darah merah. Kelaian darah ini dikenal dengan penyakit leukemia.

Produksi sel darah putih abnormal yang berlebihan pada tubuh merupakan indikasi utama penyakit kanker darah atau yang dikenal dengan penyakit leukemia (Willy, 2019). Penyakit leukemia ini sangat berbahaya. Siapa saja dapat terkena penyakit ini, tidak mengenal usia dari anak-anak hingga dewasa. Leukemia limfoblastik akut (LLA) atau *acute lymphoblastic leukemia* (ALL) merupakan salah satu jenis kanker darah. Leukemia akut terjadi bila sumsum tulang belakang berlebihan dalam

Article History:

Received: November 28th 2020; **Accepted:** March, 6th 2021

Rekayasa ISSN: 2502-5325 has been Accredited by Ristekdikti (Arjuna) Decree: No. 23/E/KPT/2019 August 8th, 2019 effective until 2023

Cite this as:

Syidada, S & Hariyanto, B. (2021). Identifikasi Acute Lymphoblastic Leukemia pada Citra Mikroskopis Menggunakan Algoritma Naïve Bayes. *Rekayasa* 14 (1). 78-83. doi: <https://doi.org/10.21107/rekayasa.v14i1.9110>

© 2021 Shofiya Shidada, Bagus Hariyanto

memproduksi sel darah putih limfosit muda atau disebut dengan limfoblast. Kasus LLA banyak ditemui pada anak-anak, akan tetapi juga dapat menyerang orang dewasa.

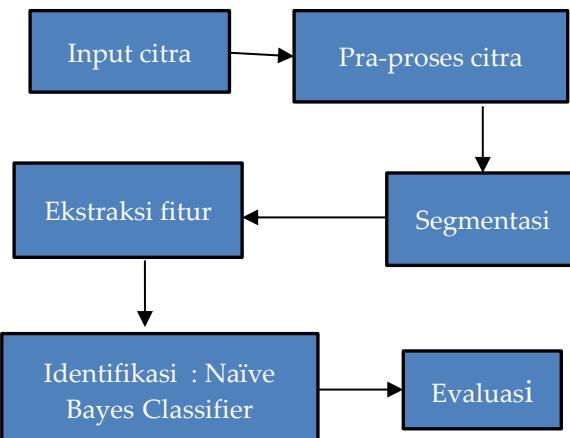
Leukimia didiagnosa dengan melakukan pemeriksaan fisik dan penunjang (Hasyimzoem, 2014). Pemeriksaan fisik dengan mengidentifikasi gejala-gejala yang dikeluhkan pasien. Sedangkan pemeriksaan penunjang adalah melakukan tes hitung darah lengkap. Tes hitung darah merupakan prosedur penghitungan banyaknya komponen sel-sel darah. Selain itu juga dilakukan pengamatan secara detail terhadap bentuk sel darah. Sedikitnya jumlah eritrosit dan bentuk sel leukosit tidak normal merupakan indikasi adanya penyakit leukemia. Pengamatan bentuk sel menggunakan mikroskop serta memerlukan keahlian khusus dari ahli hematologi (Setiawan, 2010). Selain itu proses identifikasi memakan waktu yang lama. Faktor lain yang mempengaruhi adalah tergantung pada subyektifitas pengalaman ahli hematologi, ketelitian dan konsentrasi (Putzu *et.al.*, 2014). Identifikasi sel leukemia secara otomatis membantu ahli hematologi dalam melakukan tugasnya. Identifikasi otomatis ini dilakukan berdasarkan penampakan visual dari citra mikroskopis sel darah. Hal ini dapat dilakukan dengan pengolahan citra mikroskopis sel darah (Syidada *et.al.*, 2014).

Penelitian tentang identifikasi sel leukemia secara otomatis telah banyak dilakukan (Hiremath, 2010; Reta *et.al.*, 2010; Suratin, 2015). Berbagai metode untuk klasifikasi sel leukemia digunakan oleh para peneliti seperti SVM (*Support Vector Machine*), logika fuzzy, *rule based system*, neural network (Pradana *et.al.*, 2013). *Naïve Bayes Classifier* merupakan metode klasifikasi yang menerapkan teorema bayes menggunakan perhitungan probabilitas sederhana. Namun, metode ini cukup handal dalam melakukan klasifikasi, terutama pada data numerik. Chandra, *et.al.*, (2020) menerapkan metode bayes untuk deteksi penyakit cacar air. Berdasarkan pertimbangan tersebut, pada penelitian ini penulis melakukan identifikasi sel ALL pada citra mikroskopis sel darah menggunakan metode *Naïve Bayes*.

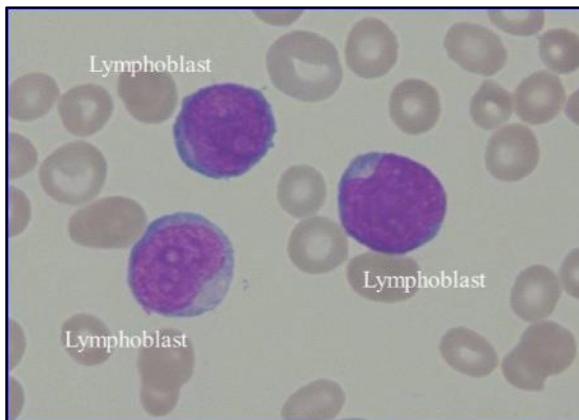
METODE PENELITIAN

Langkah-langkah pelaksanaan penelitian disajikan pada Gambar 1 yaitu diawali dengan mempersiapkan input data citra mikroskopis. Berikutnya melakukan pra-proses pada citra. Pra-proses yang dilakukan dalam penelitian ini meliputi

cropping (pemotongan citra). Tahap selanjutnya adalah segmentasi citra. Bagian citra yang disegmentasi bertujuan untuk memperoleh bagian gambar area sel leukosit. Langkah berikutnya adalah ekstraksi ciri dari hasil segmentasi. Ekstraksi ciri diperoleh dengan melakukan pengamatan terhadap karakteristik sel darah putih. Hasil ekstraksi berupa fitur atau ciri-ciri atau data atribut untuk identifikasi atau klasifikasi sel ALL.

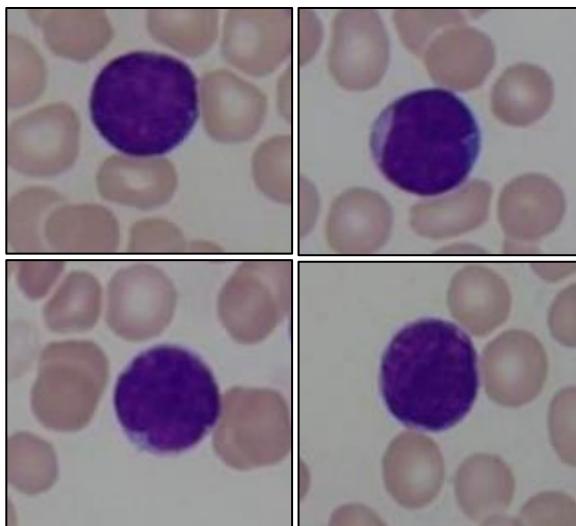


Gambar 1. Alur Pengolahan Data



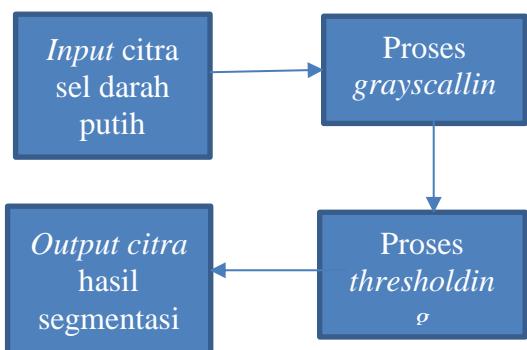
Gambar 2. Citra Mikroskopis Sel Acute Lymphoblast Leukemia

Data citra mikroskopis sel darah diperoleh dari berbagai sumber di internet. Dari data input citra ini kemudian dilakukan pra-proses citra. Pra-proses yang dilakukan pengambilan citra sel darah putih yang sudah di *cropping* secara manual oleh peneliti yang disesuaikan dengan kebutuhan sistem. Data citra mikroskopis sel darah merah sebagaimana ditunjukkan oleh Gambar 2. Selanjutnya dari gambar tersebut dilakukan pemotongan secara manual sehingga diperoleh data citra yang digunakan untuk melakukan klasifikasi. Hasil pemotongan citra mikroskopis ditunjukkan oleh Gambar 3.

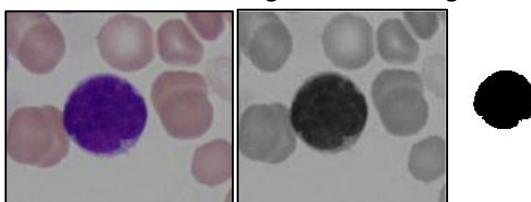


Gambar 3. Data Input Citra

Tahap Segmentasi citra bertujuan untuk memisahkan obyek-obyek yang ada pada citra (Suryani et.al., 2014; Pradana et.al., 2013). Metode segmentasi untuk memperoleh area sel darah putih menggunakan metode *thresholding*. Metode ini memanfaatkan perbedaan terang gelap dari citra grayscale (Rafael Gonzales, 2018). Langkah-langkah segmentasi citra yang dilakukan seperti pada Gambar 4. Hasil segmentasi ditunjukkan oleh Gambar 5.



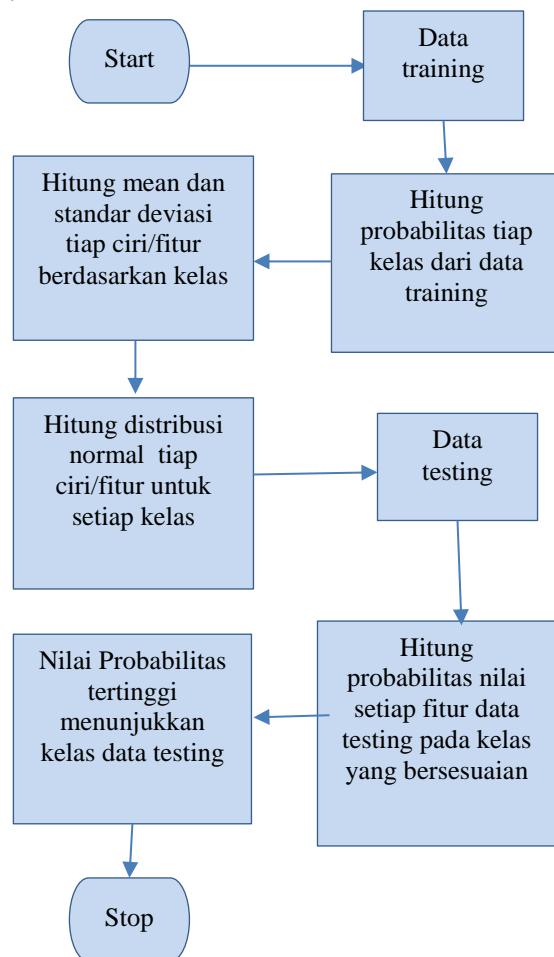
Gambar 4. Blok Diagram Proses Segmentasi



Gambar 5. Ilustrasi Proses Segmentasi

Ekstraksi ciri bertujuan untuk mengetahui karakteristik atau ciri obyek hasil segmentasi. Fitur-fitur ini diperoleh dengan melakukan perhitungan pada obyek yang telah disegmentasi. Beberapa ekstraksi ciri yang akan diekstrak yaitu luas area, *perimeter*, *compactness*, tingkat kebulatan

(*roundness*) dan *eccentricity*. Luas area adalah banyaknya jumlah piksel penyusun objek sel dara putih pada citra (Fifin, 2010). Luas area menunjukkan ukuran atau besar objek. Perimeter adalah panjang keliling dari obyek. Keliling ini dihitung dari banyaknya piksel yang berada pada tepi objek. Ciri ukuran besar kecilnya inti sel baik limfosit maupun limfoblast tidak beraturan. Perhitungan ukuran sel limfosit dilakukan dengan menggunakan Rumus (1). Ekstraksi ciri yang kedua adalah tingkat kebulatan dari sel limfoblas. Perhitungan tingkat kebulatan sel menggunakan Rumus (2). Sebagaimana yang terlihat pada Gambar 2 bahwa bentuk dari sel limfoblas mendekati bulat. Nilai *roundness* menunjukkan tingkat kebulatan dari sel.



Gambar 6. Alur Metode Naive Bayes

Tahap identifikasi citra digunakan untuk mengelompokkan sebuah data pada suatu citra ke dalam sejumlah kelas. Identifikasi atau klasifikasi dilakukan dengan menggunakan data dari ekstraksi fitur yang sudah dilakukan di tahapan sebelumnya. Proses identifikasi ini terdiri dari tahap pelatihan (*training*) dan tahap uji (*testing*). Citra mikroskopis

sel darah yang digunakan terbagi menjadi sel leukosit normal dan sel leukosit lymposit akut. Identifikasi sel darah putih ini dilakukan dengan menerapkan metode *Naïve Bayes Classifier* (NBC).

Gambar 6 menunjukkan alur dari identifikasi menggunakan Naïve bayes. Langkah pertama adalah membaca data *training*. Data fitur yang diperoleh merupakan data numerik. Selanjutnya adalah menghitung probabilitas tiap kelas dari data *training*. Tahap berikutnya menghitung nilai mean menggunakan Rumus (3) dan standar deviasi untuk setiap ciri/fitur menggunakan Rumus (4). Kemudian mengitung distribusi normal tiap ciri untuk setiap kelas menggunakan Rumus (5). Sampai pada tahap ini, proses *training* yang merupakan pembentukan model NBC menggunakan data *training* telah selesai.

$$\text{Compactness} = \frac{\text{Perimeter}^2}{\text{Area}} \dots\dots\dots(1)$$

$$\text{Roundness} = \frac{4\pi \cdot \text{area}}{\text{perimeter}^2} \dots\dots\dots(2)$$

$$\mu = \frac{\sum_{i=1}^n x_i}{n}$$

$$\mu = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n} \dots\dots\dots(3)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n-1}} \dots\dots\dots(4)$$

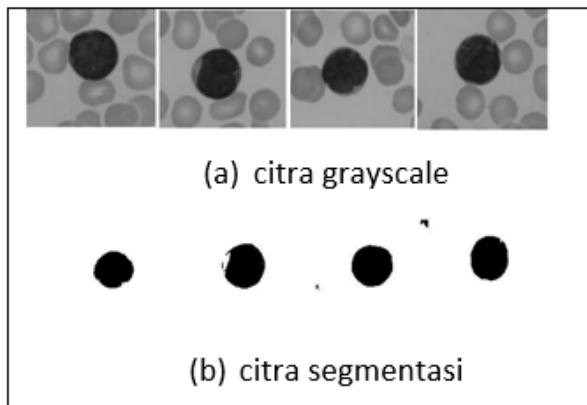
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \dots\dots\dots(5)$$

Proses klasifikasi data uji (*testing*) dilakukan dengan menghitung probabilitas setiap fitur terhadap kelas yang bersesuaian menggunakan parameter model NBC yang telah diperoleh. Selanjutnya dengan menggunakan nilai-nilai yang diperoleh sebelumnya yaitu mean, standart deviasi dan probabilitas untuk mendapatkan hasil klasifikasi. Nilai probabilitas terhadap setiap kelas dibandingkan kemudian nilai probabilitas tertinggi menentukan identifikasi data pada kelas tersebut.

HASIL DAN PEMBAHASAN

Penerapan identifikasi sel leukemia dan uji coba menggunakan pada data citra mikroskopis sel darah sebanyak 61 terdiri dari 50 citra dengan sel darah putih leukemia dan 11 citra dengan sel darah putih normal. Beberapa data citra diambil dari database citra ALL (Scotti, 2010). Hasil proses segmentasi menghasilkan pemisahan didalam citra antara objek sel darah putih dengan background, sehingga akan menghasilkan suatu citra sel darah

putih tanpa terdapat sel darah merah didalam citra tersebut. Proses segmentasi dilakukan dengan menggunakan metode *thresholding* untuk mengetahui nilai ambang suatu citra kemudian citra diubah menjadi hitam putih atau citra biner. Uji coba *thresholding* dilakukan menggunakan software matlab dimana pencarian nilai ambang dilakukan dengan otomatis dan didapatkan nilai ambang terbaik dari setiap citra sel yang diinputkan. Beberapa hasil segmentasi ditunjukkan oleh Gambar 7.



Gambar 7. Hasil Proses Segmentasi

Proses ekstraksi dilakukan untuk mendapatkan nilai nominal dari fitur luas area, perimeter, roundness, eccentricity dan compactness. Dari ekstraksi fitur ini dihasilkan sebuah data numeric. Sebagian hasil ekstraksi fitur seperti yang ditunjukkan oleh Tabel 1.

Data hasil ekstraksi fitur ini digunakan untuk melakukan identifikasi dengan menerapkan metode klasifikasi *Naïve Bayes Classifier*. Langkah pertama dalam klasifikasi adalah membagi data menjadi dua. Data training untuk membangun model pengklasifikasi. Data uji untuk mengevaluasi performa sistem pengklasifikasi. Dalam uji coba yang telah dilakukan dibagi menjadi 3 skenario, yaitu dengan membagi dataset dengan prosesntase data training dengan data testing sebanyak 80%-20%, 70%-30% dan 60%-40%.

Dari skenario uji yang pertama dengan data training 80% dari 61 sampel data dan data testing diambil 20% dari 61 sampel data secara random, didapatkan sebuah data acak dari 20% data testing sebanyak 12 data. Dari 12 data uji random tersebut dihasilkan sebuah prediksi hasil data uji benar sebanyak 9 data dan data uji salah sebanyak 3 data. Dengan memasukkan rumus akurasi $9/12 \times 100\%$ sehingga dihasilkan sebuah nilai akurasi dari 12 data uji adalah sebesar 75%.

Tabel 1. Hasil Ekstraksi Fitur

Image	Luas Area	Perimeter	Roundness	Eccentricity	Compactness
1	21962	590	0.79283	0.47544	15,8397
2	19564	556	0.79528	0.39702	15,8013
3	18980	548	0.79423	0.45516	15,8221
4	22632	600	0.79001	0.54031	15,9067
5	20368	568	0.79334	0.47204	15,8397
6	20216	566	0.793	0.48412	15,8467
7	20414	578	0.76786	0.73128	16,3654
8	21082	582	0.78212	0.64396	16,067
9	21546	590	0.77781	0.67606	16,1561
10	23496	616	0.77811	0.67087	16,1498

Dari skenario uji yang kedua dengan data training 70% dari 61 sampel data dan data testing diambil 30% dari 61 sampel data secara random, didapatkan sebuah data acak dari 30% data testing 18 data Dari 18 data uji random tersebut dihasilkan sebuah prediksi hasil data uji benar sebanyak 14 hasil uji benar dan 4 data uji salah. Dengan memasukkan rumus akurasi $14/18 \times 100\%$ sehingga dihasilkan sebuah nilai akurasi dari 18 data uji adalah sebesar 77,78%. Dari skenario uji yang ketiga dengan data testing diambil 40% dari data secara random, didapatkan sebuah data acak 25 data, didapatkan sebuah hasil data uji benar sebanyak 20 hasil uji benar dan 5 data uji salah. Dengan memasukkan rumus akurasi $20/25 \times 100\%$ sehingga dihasilkan sebuah nilai akurasi dari 25 data uji adalah sebesar 80%. Visualisasi hasil akurasi identifikasi sel leukemia menggunakan algoritma naïve bayes ditunjukkan oleh Tabel 2.

Tabel 2. Hasil Skenario Uji Coba Pertama

Skenario	Data Training	Data Testing	Akurasi
I	80%	20%	70%
II	70%	30%	77.78%
III	60%	40%	80%

KESIMPULAN DAN SARAN

Hasil eksperimen menunjukkan bahwa proses segmentasi citra menggunakan metode *thresholding* dapat memisahkan objek sel darah putih dengan *background* cukup baik dengan mencari nilai ambang terbaik. Dan dari skenario uji dalam tahap identifikasi yang dilakukan dengan tiga skenario, didapatkan sebuah hasil nilai akurasi sebesar 80% dari 25 data uji pada skenario 3

dengan proporsi perbandingan data latih dengan data uji sebesar 60%-40%. Untuk pengembangan proses identifikasi ini dapat di implementasikan ke dalam software yang berbeda, metode yang berbeda, mulai dari proses segmentasi yang tidak hanya menggunakan thresholding, dan juga proses identifikasi dapat digunakan metode lain selain metode Naïve Bayes

DAFTAR PUSTAKA

- Chandra,E.M & Yulindon, R. H. (2020). Implementasi Sistem Pakar Guna Mendiagnosa Penyakit Cacar Air Dengan Metode Bayes. *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, 10(1), 21–26.
- Fifin, D. R. (2010). Pengenalan pola citra leukosit dengan metode ekstraksi fitur citra. *Jurnal Pendidikan Fisika Indonesia*, 6, 133–137.
- Fitra, Y. A., Tjandrasa, H., & Wijaya, A. Y. (n.d.). *Implementasi segmentasi nukleus pada citra jaringan kanker dengan pendekatan berbasis morfologi*. 1–9.
- Hasyimzoem,N.C. (2014). Leukemia Limfoblastik Akut Pada Dewasa Dengan Multiple Limfadenopati. *Medula*, 2(1), 30–38.
- Hiremath, P. S. (2010). *Automated Identification and Classification of White Blood Cells (Leukocytes) in Digital Microscopic Images*.
- Noercholis, A., & Muslim, M. A. (2013). *Ekstraksi Fitur Roundness untuk Menghitung Jumlah Leukosit dalam Citra Sel Darah Ikan*. 7(1), 35–40.
- Putzu, L., Caocci, G., & Di, C. (2014). Leucocyte

- classification for leukaemia detection using image processing techniques. *Artificial Intelligence In Medicine*. <https://doi.org/10.1016/j.artmed.2014.09.002>
- Rafael Gonzales. (2018). *Digital Image Processing*.
- Reta, C., Leopoldo Altamirano, Gonzales, J. A., Diaz, R., & Guichard, J. S. (2010). Segmentation of Bone Marrow Cell Images for Morphological Classification og Acute Leukemia. *Proceedings of the Twenty-Third International Florida Artificial Intelligence Research Society Conference (FLAIRS 2010)*, (Flairs), 86–91.
- Scotti, F. (2010). Acute Lymphoblastic Leukemia Image Database for Image Processing. Diambil dari <https://homes.di.unimi.it/scotti/all/>
- Setiawan, A. (2010). Pengembangan Alat Penghitung Sel Darah. Diambil dari <https://www.dw.com/id/pengembangan-alat-penghitung-sel-darah/a-6326036>
- Suratin, M. D. (2015). Identifikasi Sel Acute Lymphoblastic Leukemia (ALL) pada Citra Peripheral Blood Smear Berdasarkan Morfologi Sel Darah Putih. *elektronik Jurnal Arus Elektro Indonesia (eJAEI)*, 7–12.
- Suryani, E., Salamah, U., & Wijaya, A. A. (2014). *Identifikasi Penyakit Acute Myeloid Leukemia (AML) Menggunakan ' Rule Based System ' Berdasarkan Morfologi Sel Darah Putih Studi Kasus: AML2 dan AML4*. 2014(November), 193–199.
- Syidada, S., Suciati, N., & Faticahah, C. (2014). Segmentasi Sel Darah Putih Berdasarkan Warna Menggunakan K-Means Dan Operasi Morfologi. *Melek IT Information Technology Journal*, 3(2).
- Pradana, T.P & Suryani, E. W. (2013). *Pemanfaatan Seed Region Growing Segmentation dan Momentum Backpropagation Neural Network untuk Klasifikasi Jenis Sel Darah Putih*.
- Willy, T. (2019). Leukemia. Diambil dari Alodokter website: <https://www.alodokter.com/leukemia>