

DETEKSI SEL DARAH SECARA OTOMATIS DENGAN *EXTREME LEARNING MACHINE*

Lina Lina¹, Jody Setiawan², Michelle Augustine³

^{1,2,3}Program Studi Teknik Informatika, Fakultas Teknologi Informasi, Universitas Tarumanagara, Jakarta, Indonesia

e-mail :¹lina@untar.ac.id, ²jody.535180064@stu.untar.ac.id, ³michelle.535200055@stu.untar.ac.id,
corresponding author: lina@untar.ac.id

ABSTRACT

Blood tests in humans play an important role not only to determine a person's blood type but also to detect diseases. Blood tests are carried out by composing blood preparations and checking them in the laboratory. With the development of technology in computer vision, the examination of blood preparations can be done through digital images of the microscope preparat. This study was conducted to detect blood cells automatically from digital images using the Extreme Learning Machine (ELM) method. The most obvious difference between ELM and other Neural Networks is in the number of its hidden layer. The ELM architecture only consists of 1 layer, while the hidden layer architecture in other Neural Networks may contain more than 1 layer. In the learning process, ELM utilizes the Moore Penrose Pseudoinverse matrix inverse theory which has the best generalization results with faster computation time. Experiments were carried out by digitizing the image of the blood cells. These images consisted of two types, the stained images and the unstained images. In the stained microscopic images, each component of the blood cells can be clearly seen and easily identified. Meanwhile, in the image of the unstained preparations, the components of the blood cells cannot be distinguished or seen directly by the human eye. The experimental results show that the developed method can achieve 95.83% accuracy for the stained image and 68.6% accuracy for the unstained image. Furthermore, the average testing accuracy of varied white blood cell types has an accuracy of 81.36%.

Keywords: Automatic detection, Blood cell detection, Extreme Learning Machine

INTISARI

Pemeriksaan darah pada manusia memegang peranan penting tidak hanya untuk mengetahui jenis golongan darah seseorang tetapi juga dalam mendeteksi penyakit. Pemeriksaan darah dilakukan dengan membuat preparat darah dan mengeceknya di laboratorium. Dengan berkembangnya teknologi dalam pengolahan citra dan visi komputer, pemeriksaan preparat darah dapat dilakukan melalui citra digital dari preparat. Studi ini dilakukan untuk mendeteksi sel darah secara otomatis dari citra digital menggunakan metode Extreme Learning Machine (ELM). Perbedaan paling nyata antara ELM dengan Neural Network yang lain berada pada hidden layer-nya, yakni pada arsitektur ELM hanya terdiri atas 1 lapisan, sedangkan arsitektur hidden layer pada Neural Network lainnya dapat lebih dari 1 lapisan. Dalam proses pembelajarannya, ELM memanfaatkan teori invers matriks Moore Penrose Pseudoinverse yang memiliki hasil generalisasi terbaik dengan waktu komputasi yang lebih cepat. Eksperimen dilakukan dengan membuat preparat sel darah, kemudian melakukan proses digitalisasi terhadap citra preparat. Citra preparat yang digunakan terdiri dari dua jenis yaitu citra yang diberi cairan pewarna Wright dan citra yang tanpa diberi cairan pewarna. Pada citra preparat yang diwarnai, setiap komponen sel darah dapat dengan jelas terlihat dan mudah diidentifikasi. Sedangkan pada citra preparat yang tidak diwarnai, maka komponen sel darah tidak dapat dibedakan dan terlihat secara langsung oleh mata manusia. Hasil eksperimen menunjukkan bahwa metode yang digunakan dapat menghasilkan akurasi validasi yang tinggi yaitu sebesar 95.83% pada citra preparat yang diwarnai dan 68.6% untuk citra preparat tanpa pewarnaan. Selanjutnya, akurasi pengujian rata-rata terhadap variasi jenis sel darah putih memiliki akurasi sebesar 81.36%.

Kata kunci: Deteksi otomatis, Deteksi sel darah, Extreme Learning Machine

1. PENDAHULUAN

Darah adalah cairan penting yang berada di dalam tubuh, darah terdiri antara plasma darah dan sel darah. Sel – sel darah ini terdiri atas tiga jenis, yaitu sel darah merah(eritrosit), sel darah putih (leukosit), dan keping darah (trombosit). Sel darah putih mempunyai peranan untuk menghasilkan antibodi sebagai pertahanan tubuh. Oleh karena itu, darah sangat penting dalam mengetahui penyakit – penyakit yang ada di dalam tubuh, banyak penyakit yang berhubungan dengan darah. Ketidakseimbangan jumlah antara sel darah merah dan sel darah putih juga dapat menyebabkan masalah Kesehatan (Huang, 2006). Sel darah putih atau leukosit merupakan salah satu sel pembentuk komponen darah yang berfungsi untuk membantu tubuh dalam melawan berbagai penyakit dan sebagai bagian dari sistem kekebalan tubuh. Mengingat pentingnya fungsi sel darah putih dalam tubuh, maka

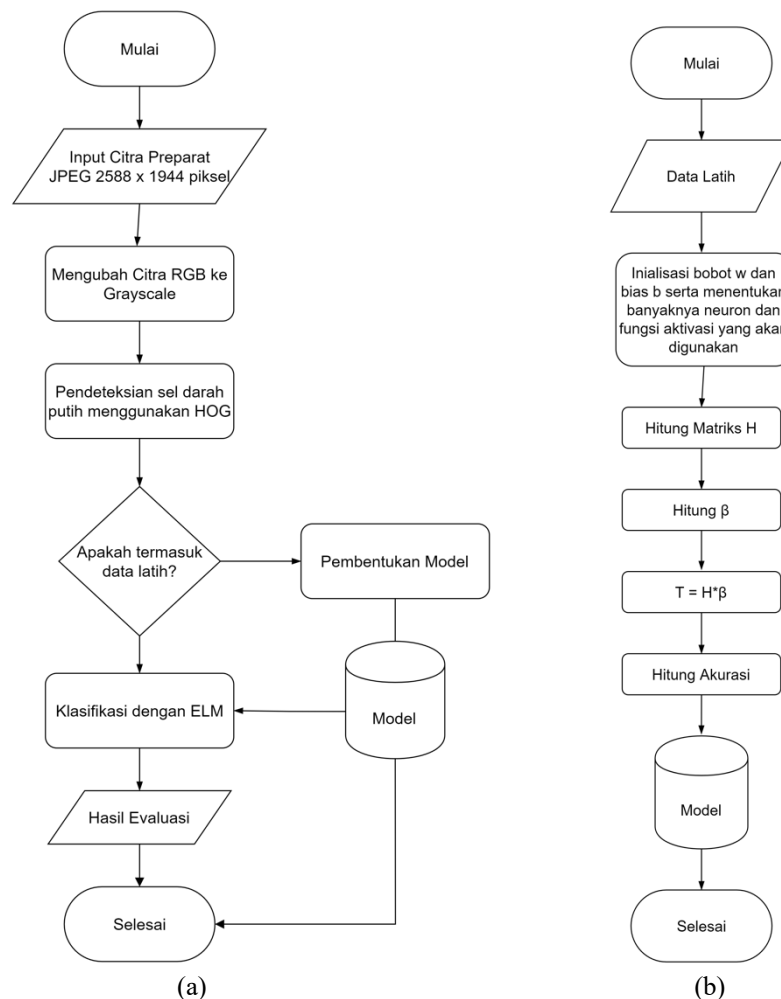
sebuah sistem berbasis teknologi yang dapat mendeteksi keberadaan sel darah putih secara otomatis sangat diperlukan.

Jaringan saraf tiruan dan metodologi *deep learning* sering diterapkan untuk melakukan berbagai analisis terkait kesehatan, seperti deteksi penyakit menular (Nugroho, 2018), deteksi sel darah (Ruberto, 2020)(Hedge, 2019)(Lina, 2016)(Lina, 2021), dan lain sebagainya. Dalam penelitian ini, dilakukan pembuatan suatu sistem untuk mendeteksi sel darah secara otomatis dari citra digital menggunakan metode *Extreme Learning Machine* (ELM). Perbedaan utama antara ELM dengan metodologi jaringan saraf buatan yang lain berada pada penggunaan lapisan tersembunyi, yakni pada arsitektur ELM hanya terdiri atas 1 lapisan, sedangkan jaringan saraf buatan lainnya memungkinkan penggunaan lebih dari 1 lapisan tersembunyi pada arsitektur lainnya. Dalam proses pembelajarannya, ELM memanfaatkan teori invers matriks Moore Penrose Pseudoinverse yang memiliki hasil generalisasi terbaik dengan waktu komputasi yang lebih cepat. Eksperimen dilakukan dengan menganalisis dua jenis preparat yaitu citra yang diberi cairan pewarna Wright dan citra yang tanpa diberi cairan pewarna. Pada citra preparat yang diwarnai, setiap komponen sel darah dapat dengan jelas terlihat dan mudah diidentifikasi. Sedangkan pada citra preparat yang tidak diwarnai, maka komponen sel darah tidak dapat dibedakan dan terlihat secara langsung oleh mata manusia. Selain itu, pendeteksian juga dilakukan terhadap berbagai jenis sel darah putih, yang mencakup basophil, eosinophil, neutrophil, limfosit dan monosit.

Struktur penulisan dari studi ini meliputi latar belakang pada bab 1, diikuti dengan penjelasan mengenai sistem deteksi yang dirancang menggunakan metode ELM pada bab 2. Bab 3 mencakup eksperimen yang dilakukan termasuk data yang digunakan. Bab 4 menyajikan kesimpulan dan saran untuk penelitian lebih lanjut.

2. SISTEM DETEKSI SEL DARAH OTOMATIS

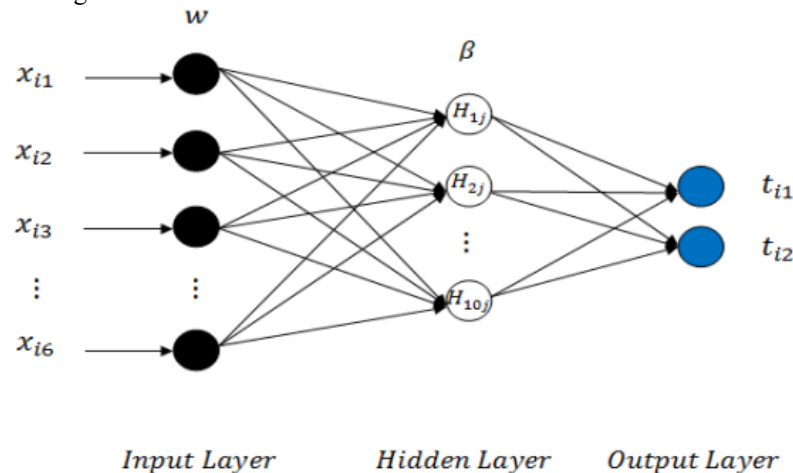
Sistem yang dirancang adalah sistem deteksi sel darah putih secara otomatis dengan metode *Extreme Learning Machine* (ELM). Sistem ini bertujuan untuk mendeteksi sel darah putih dari citra preparat dengan pewarnaan dan tanpa pewarnaan. Tahapan pada sistem yang dirancang mencakup tahap pelatihan dan tahap pengujian untuk metode ELM dapat dilihat pada Gambar 1.



Gambar 1. Tahapan sistem deteksi yang dirancang (a) skema pelatihan (b) skema pengujian

2.1 Histogram of Oriented Gradients (HOG)

Histogram of Oriented Gradients (HOG) adalah salah satu metode pengolahan citra untuk pendeteksian objek. Teknik ini menghitung nilai gradien dari suatu area tertentu dalam sebuah citra. Setiap citra memiliki ciri khas yang ditunjukkan oleh distribusi gradien. Karakteristik ini diperoleh dengan membagi citra menjadi area kecil, yang disebut sebagai “sel”. Setiap sel terdiri dari histogram gradien. Kombinasi histogram akan menjadi deskriptor yang mewakili suatu objek. Langkah pertama dari HOG adalah menghitung gradien dari citra *input*. Cara yang paling umum yang digunakan untuk menghitung gradien adalah dengan menggunakan filter Sobel secara horizontal maupun vertikal (Collobert, 2001). Selanjutnya, tetapkan bagian yang disebut sebagai sel. Setiap piksel dalam sel memiliki nilai histogramnya sendiri berdasarkan nilai yang diperoleh dari perhitungan gradien. Sel yang digunakan adalah berukuran 2x2 atau 8x8 piksel. Hasil dari fitur HOG akan diubah menjadi *feature vector* untuk diproses dengan metode ELM.



Gambar 2. Struktur dari Metode *Extreme Learning Machine*

2.2 Extreme Learning Machine (ELM)

Metode ELM merupakan suatu jaringan saraf tiruan berjenis *feedforward* dengan satu lapisan tersembunyi. ELM ini memiliki keunggulan yaitu dikarenakan hanya menggunakan satu lapisan tersembunyi, maka metode ini memiliki kecepatan proses yang tinggi. Cara kerja dari ELM adalah menerima masukan berupa *feature vector* secara acak, kemudian vektor tersebut diproses menggunakan perhitungan aktivasi *softmax* yang akan menghasilkan luaran bobot jaringan saraf tiruan berupa nilai yang mewakili jumlah kelas target. Vektor masukan disimbolkan dengan x_n , bobot w , satu lapisan layar tersembunyi H , serta kelas target t_j . Setiap indeks yang ada menentukan kelasnya sendiri. Struktur dasar dari metode ELM dapat dilihat pada Gambar 2.

Proses pelatihan pada metode ELM bertujuan untuk mendapatkan nilai bobot *output* yang kemudian digunakan untuk perhitungan dalam proses pengujian. Berikut adalah tahapan proses pelatihan dari metode ELM (Huang, 2006):

1. Inisialisasi bobot *input* dan bias secara acak. Inisialisasi nilai yang diberikan pada *input* w (bobot) dengan rentang $[-1,1]$ dan ukuran matriks j (banyak neuron tersembunyi) \times k (banyak neuron *input*). Kemudian pada nilai b (bias) dengan rentang $[0,1]$ dengan matriks $1 \times$ banyak neuron tersembunyi.
2. Hitung nilai dari:

$$H_{init} = X_{training} \cdot W^T + b \quad (1)$$

dengan H_{init} sebagai matriks awal untuk lapisan tersembunyi, $X_{training}$ sebagai *input* pelatihan, W^T sebagai matriks transpos bobot input, dan b adalah bias.

3. Hitung lapisan tersembunyi (H) menggunakan fungsi aktivasi biner Sigmoid sebagai proses perhitungan untuk hasil H menggunakan persamaan 2.

$$H = \frac{1}{1 + \exp(-H_{init})} \quad (2)$$

dengan H sebagai matriks luaran dari lapisan tersembunyi.

4. Hitung matriks *Moore-Penrose Generalized Inverse* (H^+) yang diperlukan untuk proses invers matriks. Secara umum, matriks invers adalah $n \times n$, sedangkan dalam perhitungan ini ukuran matriks adalah $n \times m$. Sehingga, perlu dilakukan perhitungan dengan menggunakan invers umum *Moore-Penrose*.

$$H^+ = (H^T \cdot H)^{-1} \cdot H^T \quad (3)$$

dengan H^+ sebagai Moore-Penrose Generalized Inverse, H^T sebagai matriks transpos.

- Langkah terakhir dalam proses pelatihan adalah menghitung bobot *output* (β) dengan matriks luaran dari lapisan tersembunyi dan Y adalah matriks target.

Dalam proses pengujian, ELM bertujuan untuk mendapatkan nilai ramalan. Berikut adalah tahapan dari ELM dalam proses pengujian:

- Menyimpan bobot *input*, bias, dan nilai dari hasil proses pelatihan.
- Hitung nilai dari:

$$H_{init} = X_{testing}.W^T + b \quad (4)$$

dengan H_{init} sebagai matriks awal untuk lapisan tersembunyi, $X_{testing}$ sebagai masukan pelatihan, W^T sebagai matriks transpos bobot input, dan b adalah bias.

- Hitung luaran lapisan tersembunyi (H) menggunakan fungsi aktivasi biner Sigmoid yang terdapat pada persamaan 2 dengan nilai masukan yang dihasilkan dari perhitungan pada persamaan 4.
- Langkah terakhir dalam proses pengujian adalah menghitung hasil ramalan.

$$Y = H.\beta \quad (5)$$

dengan Y sebagai luaran prediksi, H sebagai luaran untuk lapisan tersembunyi, dan β sebagai bobot luaran.

3. HASIL DAN PEMBAHASAN

Untuk menguji kinerja sistem, dilakukan percobaan pada basis data yang telah dikumpulkan tim peneliti. Pada penelitian ini, data citra sel darah diambil menggunakan kamera yang terhubung ke mikroskop dengan pembesaran 1000 kali yang didapat dari pembesaran lensa okuler 10 kali dan pembesaran lensa objektif 100 kali. Dari basis data yang terkumpul akan dibagi menjadi data latih yang digunakan untuk membuat model dan data uji untuk menguji hasil dari model. Data yang berhasil dikumpulkan berjumlah 473 citra untuk masing-masing citra dengan pewarnaan dan citra tanpa pewarnaan. Komposisi basis data yang dikumpulkan tertera pada Tabel 1.

Tabel 1. Data yang digunakan dalam eksperimen

Data	Citra dengan Sel Darah Putih	Citra tanpa Sel Darah Putih	Total
Latih	260	-	260
Uji	66	147	213
TOTAL			473

Pengujian dilakukan dengan menggunakan data uji yang berbeda dengan data yang digunakan untuk pelatihan untuk mengetahui tingkat keakuratan dari model yang telah dilatih pada citra yang berbeda dengan citra pelatihan. Selanjutnya untuk mengukur kinerja sistem yang dikembangkan, dilakukan perhitungan *confusion matrix*. Untuk melakukan perhitungan *confusion matrix*, evaluasi dilakukan dengan mengelompokkan hasil menjadi 4 kelompok, yaitu *True Positive* (TP), *False Positive* (FP), *True Negative* (TN) dan *False Negative* (FN). Keempat kelas ini dijelaskan sebagai berikut :

- True Positive* = citra dengan sel darah putih, terdeteksi
- False Positive* = citra tanpa sel darah putih, terdeteksi
- True Negative* = citra dengan sel darah putih, tidak terdeteksi
- False Positive* = citra tanpa sel darah putih, tidak terdeteksi

Dari nilai TP, FP, TN, dan FN yang diperoleh, dilakukan perhitungan terhadap nilai akurasi, presisi, *recall*, dan *F-1 score*. Nilai akurasi berfungsi untuk menunjukkan keakuratan dalam melakukan pengklasifikasian data dengan baik. Nilai akurasi dapat dihitung dengan persamaan sebagai berikut:

$$\text{Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (6)$$

Nilai presisi berfungsi untuk menunjukkan jumlah dalam kategori positif yang diklasifikasikan dengan benar dibagi dengan total data yang diklasifikasikan positif. Nilai presisi dapat dihitung dengan persamaan sebagai berikut:

$$\text{Presisi} = \frac{TP}{TP+FP} \times 100\% \times 100\% \quad (7)$$

Nilai *recall* berfungsi untuk menunjukkan berapa persen data kategori positif yang diklasifikasi dengan benar oleh sistem. Nilai *recall* dihitung dengan persamaan berikut:

$$Recall = \frac{TP}{TP+FN} \times 100\% \times 100\% \quad (8)$$

Nilai *F1 Score* berfungsi untuk menunjukkan perbandingan rata-rata presisi dan *recall* yang dibobotkan. Nilai *F1 Score* dapat dihitung dengan persamaan sebagai berikut :

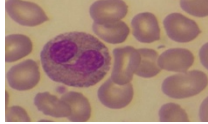
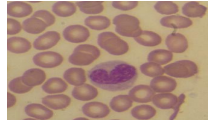
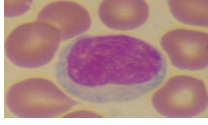
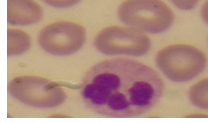
$$F1 \text{ Score} = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

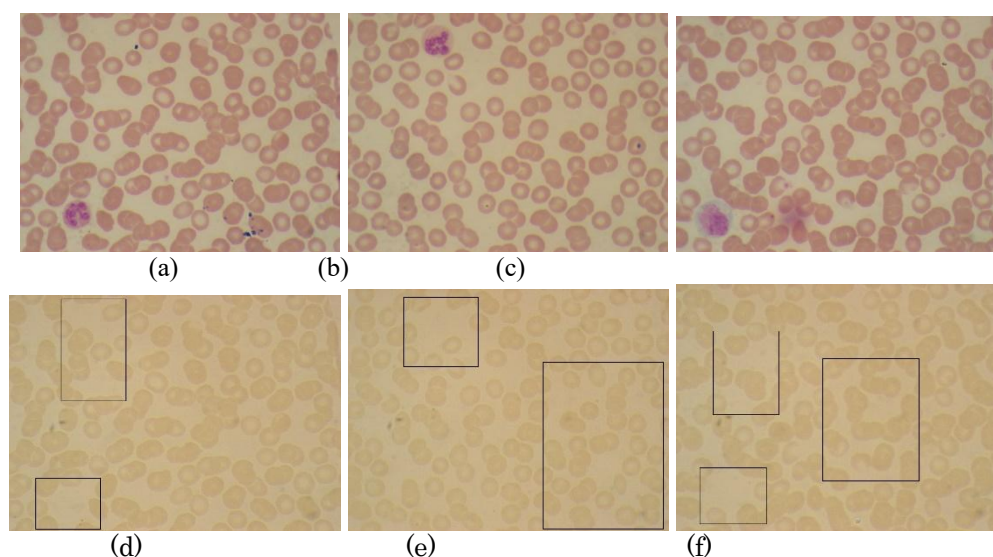
Hasil performa model dalam nilai akurasi, presisi, *recall* dan *F1 Score* dapat dilihat pada Tabel 2, sedangkan hasil pengujian dari sistem untuk setiap jenis sel darah putih tertera pada Tabel 3. Selanjutnya, contoh hasil klasifikasi citra pengujian dapat dilihat pada Gambar 3.

Tabel 2. Hasil performa model

Data	Akurasi (%)	Presisi (%)	Recall (%)	F1-score (%)
Dengan Pewarnaan	95,83	94,36	98,52	96,39
Tanpa Pewarnaan	68,6	69,95	81,59	71,15

Tabel 3. Hasil pengujian sistem untuk setiap jenis sel darah putih

Jenis Sel Darah Putih	Contoh Citra	Akurasi (%)
Eosinofil		71,42
Limfosit		84,61
Monosit		85,71
Neutrofil		83,72



Gambar 3. Contoh hasil klasifikasi citra pengujian (a-c) dengan pewarnaan, (d-f) tanpa pewarnaan

Berdasarkan hasil eksperimen, dapat terlihat bahwa sistem yang diusulkan mampu mendeteksi citra sel darah diwarnai maupun citra sel darah tanpa pewarnaan. Hasil klasifikasi tertinggi diperoleh oleh monosit dengan akurasi rekognisi adalah 85,71%, sedangkan akurasi terendah diperoleh oleh eosinofil dengan 71,42%. Monosit adalah jenis sel darah putih yang sangat mudah dideteksi oleh sistem dikarenakan jenis ini memiliki ukuran yang paling besar dibandingkan jenis sel darah putih lainnya. Kelemahan dalam sistem ini adalah bahwa sistem tidak dapat diuji dengan jenis sel basofil dikarenakan jenis sel darah putih ini tidak ditemukan dalam basis data yang dikembangkan.

4. KESIMPULAN

Berdasarkan eksperimen yang telah dilakukan untuk menguji sistem, kesimpulan yang dapat diperoleh adalah sistem yang dikembangkan mampu melakukan pendeteksian sel darah putih dari citra preparat dengan pewarnaan dengan baik. Hasil pengujian validasi terhadap klasifikasi sel darah putih memiliki akurasi validasi 95,83% untuk citra preparat dengan pewarnaan dan 68,6% untuk citra preparat tanpa pewarnaan. Hasil pengujian rata-rata terhadap variasi jenis sel darah putih memiliki akurasi sebesar 81,36%.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Direktorat Riset, Teknologi, dan Pengabdian Kepada Masyarakat, Direktorat Jenderal Pendidikan Tinggi, Riset, dan Teknologi, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi atas pendanaan penelitian yang diberikan sesuai dengan kontrak penelitian Nomor:470/LL3/AK.04/2022 dan 0779-Int-KLPPM/UNTAR/VI/2022.

DAFTAR PUSTAKA

- Collobert, R., Bengio, S., & Bengio, Y. (2001). A parallel mixture of SVMs for very large scale problems. *Advances in Neural Information Processing Systems*, 14.
- Hegde, R. B., Prasad, K., Hebbar, H., & Singh, B. M. K. (2019). Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. *Biocybernetics and Biomedical Engineering*, 39(2), 382-392.
- Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). *Extreme Learning Machine: theory and applications*. *Neurocomputing*, 70(1-3), 489-501.
- Lina, L., Chris, A., Mulyawan, B., & Dharmawan, A. B. (2016). A combination of feature selection and co-occurrence matrix methods for leukocyte recognition system. *Journal of Software Engineering and Applications*, 5(12), 101.
- Lina, L., Reynaldo, D., Danny, D., & Chris, A. (2021). White Blood Cells Detection from Unstained Microscopic Images using Modified Watershed Segmentation. *IAENG International Journal of Computer Science*, 48(4).
- Nugroho, F. (2018). *Implementasi Extreme Learning Machine Untuk Deteksi Dini Infeksi Menular Seks (IMS) Pada Puskesmas Dinoyo Kota Malang* (Doctoral dissertation, Universitas Brawijaya).
- Ruberto, C.D., Loddo, A., & Putzu, L. (2020). Detection of red and white blood cells from microscopic blood images using a region proposal approach. *Computers in biology and medicine*, 116, 103530.