

**KLASIFIKASI CITRA SEL DARAH PUTIH LIMFOBLAS DAN
NONLIMFOBLAS PADA KASUS *ACUTE LYMPHOBLASTIC LEUKEMIA*
TIPE L1 DENGAN MENGGUNAKAN METODE *K-NEAREST NEIGHBOR***

Oleh

Luthfi Muhammad Rambe

15/378784/TK/42726

Diajukan kepada Departemen Teknik Nuklir dan Teknik Fisika Fakultas Teknik
Universitas Gadjah Mada pada tanggal 25 September 2019
untuk memenuhi sebagian persyaratan untuk memperoleh derajat
Sarjana S-1 Program Studi Teknik Fisika

INTISARI

Acute Lymphoblastic Leukemia (ALL) adalah salah satu jenis leukemia akut yang kemunculannya ditandai dengan produksi sel darah putih muda secara berlebihan. ALL tipe L1 merupakan jenis ALL yang paling sering muncul (84%). Selama ini diagnosis ALL dilakukan oleh dokter/hematolog dengan menggunakan mikroskop secara manual, namun membutuhkan waktu yang lama. Meskipun mesin otomatis tersedia untuk penghitungan darah, namun daerah-daerah terpencil tidak akan mampu menyediakan alat tersebut di rumah sakit karena harga yang terlalu mahal. Oleh karena itu dilakukan penelitian dengan menggunakan mikroskop termodifikasi untuk menangkap citra sel darah putih kemudian diolah dengan pembelajaran mesin untuk menghitung serta mengklasifikasikan citra sel darah putih tersebut.

Penelitian ini menyajikan studi tentang klasifikasi citra sel darah putih pada kasus ALL tipe L1 dalam dua kelas yaitu limfoblas dan nonlimfoblas dengan menggunakan algoritma *k*-Nearest Neighbor. Fitur yang mewakili ukuran, warna, dan bentuk yang berjumlah 30 fitur berhasil diekstraksi dari citra sel darah putih yang digunakan pada penelitian.

Model pengklasifikasi terbaik didapatkan dengan menggunakan kombinasi 27 fitur, nilai parameter $k = 14$, dan rumus jarak *Euclidean distance* yang menghasilkan tingkat akurasi pengujian sebesar 89,27%. *Oversampling* data menggunakan SMOTE berhasil meningkatkan nilai *recall* limfoblas pada data uji sebesar 7,76%. Pembuangan nilai fitur yang *outlier* mampu meningkatkan akurasi pengujian sebesar 0,77% dan *F1-score* pengujian sebesar 2,81%. Hasil tersebut sudah cukup signifikan untuk mendukung keputusan dokter dalam melakukan diagnosis terhadap penyakit ALL.

Kata kunci: ALL, Sel Darah Putih, *k*-Nearest Neighbor, Fitur, Akurasi

Pembimbing Utama : Nopriadi, S.T., M.Sc., Ph.D.

Pembimbing Pendamping : Ir. Agus Arif, M.T.

**WHITE BLOOD CELL CLASSIFICATION OF LYMPHOBLAST AND
NON-LYMPHOBLAST IN ACUTE LYMPHOBLASTIC LEUKEMIA
TYPE L-1 USING K-NEAREST NEIGHBOR METHOD**

by

Luthfi Muhammad Rambe

15/378784/TK/42726

Submitted to the Department of Nuclear Engineering and Engineering Physics
Faculty of Engineering Universitas Gadjah Mada on September 25th, 2019
in partial fulfillment of the requirement for the Degree of
Bachelor of Engineering in Engineering Physics

ABSTRACT

Acute Lymphoblastic Leukemia is a type of cancer that contributes to the production of abnormal white blood cells (WBC). ALL type L1 is an ALL type that often appears (84%). Current diagnosis practiced by hematologists using a microscope which is conducted manually. However, manual observation is exhausting and time consuming. Even though the automated machines are available for blood count, certain countries will not be able to afford to equip their hospitals with the machine because the price is too expensive. This research was conducted using a modified microscope to capture the white blood cells image and then processed with machine learning to calculate and classify it.

This research presents the study on white blood cells classifying in ALL type L1 into two categories which are lymphoblasts and non lymphoblasts by using k-Nearest Neighbor algorithm. Features that represent size, color-based, and shape totaling 30 features were extracted from acute leukemia blood images used in this study.

Result show that by having $k = 14$, using Euclidean distance as distance metric, and using 27 features, the testing accuracy is 89,27%. Data oversampling using SMOTe succeeded in increasing the recall value of lymphoblast in the test data by 7.76%. The removal of the outlier feature values can increase the value of the testing accuracy by 0,77% and the testing F1-score by 2,81%. These results were significant for doctors taking a diagnosis of ALL disease.

Keywords: ALL, White Blood Cells, k-Nearest Neighbor, Feature, Accuracy

Supervisor : Nopriadi, S.T., M.Sc., Ph.D.

Co-supervisor : Ir. Agus Arif, M.T.