

# **Klasifikasi Citra Sel Darah Putih Kanker dan Non-kanker Berdasarkan Ciri Morfologi Inti Sel menggunakan Algoritma *K-Nearest Neighbors* (K-NN)**

**Wildan Mahad Tahtadi**

## **ABSTRAK**

Penyakit leukemia merupakan salah satu jenis kanker yang menyerang sel darah putih manusia. Salah satu jenis leukemia adalah *acute lymphoblastic leukemia* (ALL). Perlu penanganan yang tepat dan cepat terhadap penderita ALL. Pemeriksaan secara manual memerlukan waktu yang lama dan rentan terhadap *human error*. Untuk mengatasi hal tersebut, diperlukan aplikasi yang mampu mengidentifikasi apakah sel tersebut merupakan sel kanker atau non-kanker. Penelitian ini bertujuan untuk mengetahui akurasi dari klasifikator menggunakan algoritma K-NN berdasarkan ciri morfologi inti sel dari sel darah putih. Data citra sel darah putih akan melalui tahapan praproses, segmentasi citra, ekstraksi ciri, ekstraksi *region of interest* (RoI) dan klasifikasi yang dibagi menjadi dua fase. Fase pertama meliputi praproses, segmentasi, dan ekstraksi RoI pada citra, fase kedua meliputi ekstraksi ciri dan klasifikasi. Praproses dilakukan dengan mengubah citra yang awalnya memiliki model warna RGB ke dalam HSV, lalu diterapkan *linear contrast stretch* pada citra HSV tersebut. Segmentasi citra menggunakan *fuzzy c-means* dengan kombinasi warna citra saturasi dan *value* dengan jumlah *cluster* = 3. Hasilnya citra terbagi kedalam *cluster background*, inti sel, serta sel darah merah dan sitoplasma (dalam satu *cluster*). Setelah itu inti sel akan ekstrak dari citra dan akan dievaluasi untuk digunakan pada fase kedua. Hasil dari fase pertama berupa inti sel yang terekstrak dengan baik yaitu 247 citra dari 260 citra. Citra inti sel tersebut digunakan sebagai masukan untuk fase kedua. Pada tahap ekstraksi ciri dilakukan dengan mengambil fitur bentuk dari inti sel tersebut dan data ekstraksi ciri digunakan untuk klasifikasi dengan K-NN. Setelah dilakukan uji akurasi ,dengan data latih 80% dan data uji 20% menghasilkan akurasi yang paling baik dengan  $k = 3$  dengan presentase akurasi terhadap data latih sebesar 84.77% dan sebesar 80.00% terhadap data uji.

**Kata Kunci** : *acute lymphoblastic leukemia, klasifikasi, K-NN.*

**Klasifikasi Citra Sel Darah Putih Kanker dan Non-kanker  
Berdasarkan Ciri Morfologi Inti Sel menggunakan Algoritma *K-Nearest Neighbors* (K-NN)**

**Wildan Mahad Tahtadi**

**ABSTRACT**

Leukemia is one type of cancer that attacks human white blood cells. One type of leukemia is acute lymphoblastic leukemia (ALL). Need proper and fast handling of ALL patients. Manual examination takes a long time and is vulnerable to human error. To overcome this, an application is needed that is able to identify whether these cells are cancerous or non-cancerous cells. This study aims to determine the accuracy of the classification using the K-NN algorithm based on the morphological characteristics of the cell nucleus of white blood cells. White blood cell image data will go through preprocessing, image segmentation, feature extraction, region of interest (RoI) extraction and classification which are divided into two phases. The first phase includes preprocessing, segmentation, and RoI extraction in the image, the second phase includes extraction of features and classification. Preprocessing is done by changing the image that originally had an RGB color model into HSV, then applied linear contrast stretch to the HSV image. Image segmentation using fuzzy c-means with a combination of color image saturation and value with the number of clusters = 3. The results of the image are divided into background clusters, cell nuclei, and red blood cells and cytoplasm (in one cluster). After that the cell nucleus will extract from the image and will be evaluated for use in the second phase. The results of the first phase are cell nuclei which are well extracted, namely 247 images from 260 images. The nucleus of white blood cell image is used as input for the second phase. At the feature extraction stage is done by taking the form features of the cell nucleus and feature extraction data used for classification with K-NN. After testing the accuracy, with 80% training data and 20% test data produces the best accuracy of other data compositions with the most optimal of  $k = 3$  with the percentage of accuracy of the training data amounting to 84.77% and 80.00% of the test data.

**Kata Kunci :** *acute lymphoblastic leukemia, classification, K-NN.*