

ABSTRAK

Acute Lymphoblastic Leukemia (ALL) merupakan salah satu jenis kanker darah yang sering terjadi pada anak-anak yang berhubungan dengan infiltrasi organ lain termasuk sumsum tulang yang terkena akibat proliferasi sel limfoblastik merupakan anemia yang terlihat dengan gejala klinis pucat, lelah, takikardia, dispnea, terkadang disertai gagal jantung kongestif, ulserasi mukosa bukal, dan infeksi. Pada tahun 2022 diprediksi terjadinya kasus baru leukemia sekitar 60.650 kasus dan 24.000 jiwa meninggal dunia akibat leukemia di Amerika Serikat. Dalam mendiagnosis leukemia, terdapat beberapa cara seperti menghitung banyaknya sel darah dalam darah, kemudian untuk melihat apakah sel-sel termasuk normal atau abnormal dengan melakukan tes mikroskopik. Berdasarkan hasil tes-tes tersebut ahli kesehatan atau dokter dapat mendiagnosa leukemia, apakah termasuk dalam leukemia akut atau leukemia kronis yang membutuhkan waktu relatif lama. Pada citra mikroskopis sel darah putih pada penyakit ALL dilakukan klasifikasi untuk mengetahui tingkat akurasi yang optimal dari sebuah model. Proses klasifikasi dibagi menjadi proses pelatihan dan pengujian dilanjutkan dengan evaluasi. Seluruh data dibagi menjadi dua bagian yaitu data pelatihan dan data pengujian masing-masing 80% dan 20%. Data pelatihan dilakukan evaluasi kinerja sistem dengan membagi menjadi data *training* dan data *validation* menggunakan *k-fold cross validation*. Selanjutnya, data dilakukan pelatihan menggunakan model *ResNet-18* dengan *hyperparameter learning rate* 0.1, 0.001, dan 0.0001 serta *epoch* 60, 80, dan 100. Berdasarkan *hyperparameter* tersebut dipilih model terbaik dengan *hyperparameter epoch* 60 dan *learning rate* 0.0001 yang dilakukan pengujian terhadap data pengujian menghasilkan *accuracy* mencapai 98.07%. Sedangkan nilai *precision*, *recall*, dan $F1_{score}$ masing-masing 98.33%, 97.82%, dan 98.04% dengan waktu komputasi pengujian selama 31.78s.

Kata Kunci : Klasifikasi, *Acute Lymphoblastic Leukemia* (ALL), *Convolutional Neural Network* (CNN), *Hyperparameter*, *ResNet-18*

ABSTRACT

Acute Lymphoblastic Leukemia (ALL) is a type of blood cancer that often occurs in children associated with infiltration of other organs including bone marrow affected by lymphoblastic cell proliferation is an anemia seen with clinical symptoms of pallor, fatigue, tachycardia, dyspnea, sometimes failure accompanied by congestive heart failure, buccal mucosal ulceration, and infection. In 2022, it is predicted that there will be around 60,650 new cases of leukemia and 24,000 deaths from leukemia in the United States. In diagnosing leukemia, there are several ways such as counting the number of blood cells in the blood, then to see if the cells are normal or abnormal by doing a microscopic test. Based on the results of these tests, a health professionals or doctors can diagnose leukemia, whether it belongs to acute leukemia or chronic leukemia which takes a relatively long time. In microscopic images of white blood cells in ALL disease, classification is used to identify the optimal level of accuracy of a model. The classification process is divided into a training process and testing is followed by evaluation. All data is divided into two parts, namely training data and test data of 80% and 20%, respectively. Training data are evaluated by dividing them into training data and validation data using k-fold cross validation. Furthermore, data were trained using the ResNet-18 model with hyperparameter learning rates of 0.1, 0.001, and 0.0001 and epochs of 60, 80, and 100. Based on these hyperparameters, the best model with hyperparameter epoch 60 and a learning rate of 0.0001 was selected, which was tested on test data, resulting in an accuracy of up to 98.07%. Meanwhile, the precision, recall, and $F1_{score}$ values are 98.33%, 97.82%, and 98.04% respectively with a test computation time of 31.78s.

Keywords : *Classification, Acute Lymphoblastic Leukemia (ALL), Convolutional Neural Network (CNN), Hyperparameter, ResNet-18*

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN ORSINALITAS.....	iv
KATA PENGANTAR.....	v
ABSTRAK.....	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
DAFTAR KODE PROGRAM	xiv
DAFTAR LAMPIRAN	xv
DAFTAR SIMBOL MATEMATIKA	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.2.1 Permasalahan.....	3
1.2.2 Solusi Permasalahan	3
1.2.3 Pertanyaan Penelitian.....	3
1.3 Batasan Masalah	4
1.4 Tujuan dan Manfaat	4
1.4.1 Tujuan	4
1.4.2 Manfaat	4
1.5 Sistematika Buku Skripsi	4
BAB II KAJIAN PUSTAKA	7
2.1 Klasifikasi Citra	7

2.2	<i>Convolutional Neural Network (CNN)</i>	7
2.2.1	<i>Convolutional Layer (Lapisan Konvolusi)</i>	8
2.2.2	<i>Rectified Linear Unit (ReLU)</i>	10
2.2.3	<i>Pooling Layer (Lapisan Penggabungan)</i>	11
2.2.4	<i>Fully Connected Layer (FCL)</i>	12
2.3	<i>Residual Network (ResNet)</i>	15
2.4	<i>Arsitektur ResNet-18</i>	17
2.5	<i>Keras</i>	18
2.6	<i>Optimizer</i>	18
2.6.1	<i>Adaptive Moment Estimation (ADAM)</i>	19
2.7	<i>Evaluasi Kinerja Sistem</i>	20
2.7.1	<i>K-Fold Cross Validation</i>	20
2.7.2	<i>Matrik Evaluasi (Confusion Matrix)</i>	21
2.8	<i>Penelitian Terkait</i>	23
BAB III METODE USULAN		32
3.1.	<i>Dataset</i>	32
3.2.	<i>Arsitektur Sistem Training dan Testing</i>	32
3.3.	<i>Perhitungan Proses Pelatihan Model</i>	32
3.3.1	<i>Citra Input</i>	32
3.3.2	<i>Proses Training Data</i>	34
3.3.3	<i>Proses Testing Data</i>	42
3.4.	<i>Skenario Uji Coba</i>	48
BAB IV HASIL DAN PEMBAHASAN.....		50
4.1	<i>Lingkungan Uji Coba</i>	50
4.2	<i>Tahap - Tahap Pembuatan Program</i>	51
4.2.1	<i>Memuat Path Citra Dataset</i>	51
4.2.2	<i>Menyimpan Path Dataset Dalam File Excel</i>	53

4.2.3 Pembagian <i>Dataset</i> Pelatihan dan Pengujian.....	54
4.2.4 Augmentasi Citra	55
4.2.5 Menggabungkan <i>Path</i> Data Pelatihan dan Data Pengujian Menjadi File <i>Excel</i>	56
4.2.6 Pembagian <i>Dataset Training</i> dan <i>Validation (k-fold Cross Validation)</i>	57
4.2.7 Pembuatan Model <i>ResNet-18</i>	59
4.2.8 Proses Pelatihan (<i>Training</i>)	62
4.2.9 Proses Pengujian Model Terbaik (<i>Testing Best Model</i>)	65
4.3 Hasil Uji Coba	68
4.3.1 Hasil Uji Coba Skenario 1	68
4.3.2 Hasil Uji Coba Skenario 2	69
4.3.3 Hasil Uji Coba Skenario 3	71
4.3.4 Hasil Uji Coba Skenario 4	72
4.3.5 Hasil Uji Coba Data Pengujian pada Model Terbaik	74
4.4 Analisa Hasil Uji Coba	76
BAB V PENUTUP.....	78
5.1 Kesimpulan.....	78
5.2 Saran	78
REFERENSI.....	75
LAMPIRAN-LAMPIRAN	79
RIWAYAT HIDUP.....	120

DAFTAR GAMBAR

Gambar 2. 1 Arsitektur Dasar <i>Convolutional Neural Network</i> (CNN).....	8
Gambar 2. 2 Operasi <i>Zero Padding</i>	10
Gambar 2. 3 <i>Stride</i>	9
Gambar 2. 4 <i>Rectified Linear Unit</i> (ReLU).....	11
Gambar 2. 5 <i>Max-Pooling</i>	12
Gambar 2. 6 <i>Fully Connected Layer</i> (FCL)	13
Gambar 2. 7 <i>Residual Unit</i>	16
Gambar 2. 8 Arsitektur <i>ResNet-18</i> [25]	18
Gambar 2. 9 <i>K-Fold Cross Validation</i>	21
Gambar 2. 10 <i>Confusion Matrix</i>	21
Gambar 3. 1 Sel Sehat ALL-IDB.....	32
Gambar 3. 2 <i>Acute Lymphoblastic Leukemia</i> ALL-IDB.....	32
Gambar 3. 3 Arsitektur Sistem <i>Training</i> dan <i>Testing</i>	31
Gambar 4. 1 Grafik Visualisasi Hasil Uji Coba Skenario 1	69
Gambar 4. 2 Grafik Visualisasi Hasil Uji Coba Skenario 2	70
Gambar 4. 3 Grafik Visualisasi Hasil Uji Coba Skenario 3	72
Gambar 4. 4 Grafik Visualisasi Hasil Uji Coba Skenario 4	74

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait.....	26
Tabel 3. 1 Sampel Citra <i>Input</i>	33
Tabel 3. 2 Sampel Citra <i>Input</i> Setelah <i>Zero-Padding</i>	34
Tabel 3. 3 Matriks <i>Filter</i> Citra.....	35
Tabel 3. 4 Matriks <i>Output</i> Konvolusi	36
Tabel 3. 5 Matriks Hasil Aktivasi <i>ReLu</i>	36
Tabel 3. 6 Matriks Hasil <i>Max-Pooling Layer</i>	37
Tabel 3. 7 Hasil <i>Residual Block</i>	37
Tabel 3. 8 Hasil <i>Operasi Add</i>	37
Tabel 3. 9 Hasil <i>Operasi Add</i> Setelah Aktivasi <i>ReLu</i>	38
Tabel 3. 10 Hasil <i>Average Pooling</i>	38
Tabel 3. 11 Hasil <i>Flatten</i>	38
Tabel 3. 12 Inisialisasi Nilai Bobot dan Bias	39
Tabel 3. 13 Hasil Momen Pertama	40
Tabel 3. 14 Hasil Momen Kedua	40
Tabel 3. 15 Hasil <i>Bias Correction</i> Momen Pertama.....	41
Tabel 3. 16 Hasil <i>Bias Correction</i> Momen Kedua.....	41
Tabel 3. 17 Hasil Update Bobot dan Bias Baru <i>ADAM</i>	42
Tabel 3. 18 Sampel Citra <i>Input Testing</i> Setelah <i>Zero-Padding</i>	42
Tabel 3. 19 Matriks <i>Filter</i> Citra <i>Testing</i>	44
Tabel 3. 20 Matriks <i>Output</i> Konvolusi Citra <i>Testing</i>	44
Tabel 3. 21 Matriks Hasil Aktivasi <i>ReLu</i> Citra <i>Testing</i>	45
Tabel 3. 22 Matriks Hasil <i>Max-Pooling Layer</i> Citra <i>Testing</i>	45
Tabel 3. 23 Hasil <i>Residual Block</i> Citra <i>Testing</i>	45
Tabel 3. 24 Hasil <i>Operasi Add</i> Citra <i>Testing</i>	46
Tabel 3. 25 Hasil <i>Operasi Add</i> Setelah Aktivasi <i>ReLu</i> Citra <i>Testing</i>	46
Tabel 3. 26 Hasil <i>Average Pooling</i> Citra <i>Testing</i>	46
Tabel 3. 27 Hasil <i>Flatten</i> Citra <i>Testing</i>	47
Tabel 3. 28 Inisialisasi Nilai Bobot dan Bias Citra <i>Testing</i>	47
Tabel 3. 29 Skenario Uji Coba	49
Tabel 4. 1 Spesifikasi Perangkat <i>Keras</i>	50
Tabel 4. 2 Spesifikasi <i>Colaboratory Notebooks</i>	50

Tabel 4. 3 Spesifikasi Perangkat Lunak	50
Tabel 4. 4 Hasil Skenario 1A, 1B, dan 1C	68
Tabel 4. 5 Hasil Skenario 2A, 2B, dan 2C	70
Tabel 4. 6 Hasil Skenario 3A, 3B, dan 3C	71
Tabel 4. 7 Hasil Skenario 4A, 4B, dan 4C	72
Tabel 4. 8 Model Terbaik Berdasarkan <i>Validation Accuracy</i> dan Waktu Komputasi Pelatihan Menggunakan $k=5$	74
Tabel 4. 9 Model Terbaik Berdasarkan <i>Validation Accuracy</i> dan Waktu Komputasi Pelatihan Menggunakan $k=10$	75
Tabel 4. 10 Hasil Evaluasi Model Terbaik Menggunakan $k=5$	75
Tabel 4. 11 Hasil Evaluasi Model Terbaik Menggunakan $k=10$	76

DAFTAR KODE PROGRAM

Kode Program 4. 1 Memuat <i>Path</i> Citra <i>Dataset</i>	52
Kode Program 4. 2 Menyimpan <i>Path Dataset</i> Dalam File <i>Excel</i>	53
Kode Program 4. 3 Pembagian <i>Dataset</i> Pelatihan dan Pengujian	54
Kode Program 4. 4 Augmentasi Citra	56
Kode Program 4. 5 Pembagian <i>Dataset Training</i> dan <i>Validation (k-fold Cross Validation)</i> .	58
Kode Program 4. 6 Pembuatan Model <i>ResNet-18</i>	61
Kode Program 4. 7 Proses Pelatihan (<i>Training</i>).....	64
Kode Program 4. 8 Proses Pengujian Model Terbaik (<i>Testing Best Model</i>).....	68

DAFTAR LAMPIRAN

Lampiran 1 Hasil Uji Coba 2A	79
Lampiran 2 Hasil Uji Coba 3A	86
Lampiran 3 Hasil Uji Coba 3B	94
Lampiran 4 Hasil Uji Coba 4C	104

DAFTAR SIMBOL MATEMATIKA

Simbol	Keterangan	Pemakaian Pertama Kali Pada Halaman
*	Operasi Konvolusi	8
$\sum$	<i>Sigma Sum</i>	8
+	Tambah; Menambah; Positif	8
-	Kurang; Mengurangi; Negatif	8
=	Sama Dengan	8
$\lfloor \rfloor$	<i>Floor</i>	8
()	Kurung Biasa	8
$\leq$	Kurang Dari atau Sama Dengan	8
$>$	Lebih Dari	11
$<$	Kurang Dari	15
$\geq$	Lebih Dari atau Sama Dengan	15
m	<i>First Momentum</i>	19
v	<i>Second Momentum</i>	
t	<i>time</i>	19
β	Konstanta <i>Beta</i>	19
$\hat{m}$	<i>New First Momentum</i>	19
$\hat{v}$	<i>New Second Momentum</i>	19
g	<i>Gradient</i>	19
w	<i>Weight</i>	20
σ	<i>Learning Rate</i>	20
ϵ	<i>Konstanta Epsilon</i>	20
$\sqrt{}$	<i>Radical Sign</i>	20