

KLASIFIKASI CITRA ACUTE LYMPHOBLASTIC LEUKEMIA MENGUNAKAN ARSITEKTUR RESNET-50

Sandi Indika Saputra
(170411100052)

Dosen Pembimbing

Prof. Dr. Arif Muntasa, S.Si., M.T.
Dr. Rima Tri Wahyuningrum, S.T., M.T.

ABSTRAK

Kanker anak telah menjadi momok kesehatan masyarakat global dengan kejadian yang meningkat setiap tahunnya. Kanker anak yang paling sering terjadi adalah *Acute Lymphoblastic Leukemia* (ALL) dan memiliki jumlah kasus kanker tertinggi pada anak-anak di Indonesia. Pengamatan bentuk sel dalam pemeriksaan darah untuk mendiagnosis ALL dilakukan oleh hematolog dengan menggunakan mikroskop. Proses identifikasi yang dilakukan memakan waktu yang cukup lama, serta tergantung pada subjektif pengetahuan dari ahli hematologi, konsentrasi, dan ketelitian. Oleh sebab itu diperlukan suatu sistem yang dapat melakukan identifikasi secara otomatis pada citra sel darah putih dalam menentukan apakah terdapat penyakit ALL atau tidak didalamnya. Penelitian ini dilakukan untuk merancang sistem klasifikasi citra ALL berdasarkan citra mikroskopis sel darah putih. *Dataset* yang digunakan adalah C-NMC *Dataset* berupa sel B-limfoblas leukemia (sel kanker) dan B-limfoid normal (sel normal) dari gambar mikroskopis sediaan apusan darah tepi dengan total 12.528 gambar. Metode klasifikasi yang diterapkan adalah metode *Deep Learning* yang berbasis *Residual Network* yaitu model ResNet-50. Dilakukan pendekatan secara obyektif dengan mencari nilai yang cocok untuk *hyperparameter* model seperti *learning rate* dan *epoch*. Hasil akurasi maksimal yang didapatkan pada uji coba yang dilakukan mendapatkan performa berdasarkan akurasi prediksi sebesar 93,58% dengan nilai *epoch* yang digunakan adalah 15 dan *learning rate* 10^{-4} pada data yang tidak seimbang. Sedangkan pada data seimbang

didapatkan model yang tidak mengalami konvergen atau *overfitting* dengan hasil akurasi sebesar 89,66% dengan nilai parameter *learning rate* dan *epoch* yang sama.

Kata kunci: *Acute Lymphoblastic Leukemia, Deep Learning, Residual Network, ResNet-50.*

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN ORISINALITAS	iii
ABSTRAK	vi
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xiii
DAFTAR TABEL	xv
DAFTAR KODE PROGRAM	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.2.1 Permasalahan.....	3
1.2.2 Metode Penelitian	4
1.2.3 Pertanyaan Penelitian.....	4
1.3 Tujuan dan Manfaat Penelitian.....	4
1.3.1 Tujuan Penelitian.....	4
1.3.1 Manfaat Penelitian	4
1.4 Batasan Masalah	5
1.5 Sistematika Penulisan	5
BAB II KAJIAN PUSTAKA	7
2.1 <i>Acute Lymphoblastic Leukemia</i>	7
2.2 Data Tidak Seimbang.....	7
2.3 <i>Convolutional Neural Network</i>	9
2.2.1 <i>Convolution Layer</i>	10
2.2.2 <i>Pooling Layer</i>	13
2.2.3 <i>Fully Connected Layer</i>	14
2.4 ResNet-50.....	16
2.5 <i>Optimizer</i>	20

2.6	<i>Transfer Learning</i>	22
2.7	<i>Cross Validation</i>	23
2.8	<i>Model Summary</i>	24
2.9	<i>Area Under Receiver Operating Characteristic Curve (AUC-ROC)</i>	25
2.10	Evaluasi Sistem.....	26
2.11	Penelitian Terkait	28
BAB III METODE PENELITIAN		33
3.1	<i>Dataset</i>	33
3.2	Rancangan Sistem.....	36
3.3	Skenario Uji Coba.....	38
BAB IV HASIL DAN PEMBAHASAN		40
4.1	Lingkungan Uji Coba.....	40
4.2	Data Citra	40
4.3	<i>Undersampling Data</i>	45
4.4	Praproses Data	47
4.5	<i>Splitting Data</i>	48
4.6	Model <i>Transfer Learning</i> ResNet-50	49
4.7	<i>Training Model</i>	51
4.8	<i>History Training</i>	56
4.9	Proses Pengujian.....	58
4.10	Melakukan Prediksi Menggunakan Model.....	60
4.11	<i>Confusion Matric, Classification Report, dan ROC Curve</i>	62
4.12	Hasil Analisis Uji Coba	64
4.9.1	Uji Coba Perbandingan Bobot Model.....	65
4.9.2	Uji Coba Parameter <i>Learning Rate</i>	65
4.9.3	Uji Coba Parameter <i>Epoch</i>	77
BAB V PENUTUP		88
5.1	Kesimpulan.....	88
5.2	Saran	88
DAFTAR RUJUKAN		90
LAMPIRAN		95

DAFTAR GAMBAR

Gambar 2.1 Arsitektur CNN	10
Gambar 2.2 Operasi <i>convolution</i> dengan <i>stride</i> 2	11
Gambar 2.3 Ilustrasi <i>zero padding</i>	12
Gambar 2.4 Ilustrasi fungsi aktivasi ReLU	13
Gambar 2.5 Ilustrasi operasi <i>max pooling</i>	14
Gambar 2.6 Contoh arsitektur jaringan ImageNet.....	17
Gambar 2.7 Diagram blok arsitektur ResNet-50 [12].....	18
Gambar 2.8 Arsitektur untuk ImageNet dengan blok sesuai <i>bracket</i> nomor blok. <i>Down-sampling</i> terdapat pada conv3_1, conv4_1, dan conv5_1 dengan jumlah 2 <i>stride</i> [23]	20
Gambar 2.9 Ilustrasi Proses Model Transfer.....	23
Gambar 2.10 Ilustrasi <i>K-fold cross validation</i>	23
Gambar 2.11 Contoh ROC <i>curve</i>	25
Gambar 3.1 Alur penelitian.....	33
Gambar 3.2 Data citra pada C-NMC 2019 <i>Dataset</i>	34
Gambar 3.3 Contoh data citra sel ALL dan sel normal	35
Gambar 3.4 Rancangan sistem	36
Gambar 4.1 <i>Dataframe</i> yang memuat data citra.....	44
Gambar 4.2 <i>History Training Model Fold</i> ke-1	69
Gambar 4.3 <i>Confusion Matrix Model Fold</i> ke-1.....	70
Gambar 4.4 <i>Classification Report Model Fold</i> ke-1	70
Gambar 4.5 ROC <i>Curve Model Fold</i> ke-1.....	71
Gambar 4.6 <i>History Training Model Fold</i> ke-4	74
Gambar 4.7 <i>Confusion Matrix Model Fold</i> ke-4.....	75
Gambar 4.8 <i>Classification Report Model Fold</i> ke-4	75
Gambar 4.9 ROC <i>Curve Model Fold</i> ke-4.....	76
Gambar 4.10 <i>History Training Model Fold</i> ke-2	80
Gambar 4.11 <i>Confusion Matrix Model Fold</i> ke-2.....	80
Gambar 4.12 <i>Classification Report Model Fold</i> ke-2	81
Gambar 4.13 ROC <i>Curve Model Fold</i> ke-2.....	81

Gambar 4.14 <i>History Training Model Fold</i> ke-1	84
Gambar 4.15 <i>Confusion Matrix Model Fold</i> ke-1	85
Gambar 4.16 <i>Classification Report Model Fold</i> ke-1	85
Gambar 4.17 <i>ROC Curve Model Fold</i> ke-1	86

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait	30
Tabel 3.1 Detail <i>dataset</i> C-NMC 2019	34
Tabel 3.2 Skenario uji coba	39
Tabel 4.1 Spesifikasi Perangkat Keras dan Perangkat Lunak	40
Tabel 4.2 Hasil pelatihan model dengan uji coba <i>learning rate</i> pada data tidak seimbang	66
Tabel 4.3 Hasil pengujian model dengan uji coba <i>learning rate</i> pada data tidak seimbang	67
Tabel 4.4 Hasil pelatihan model dengan uji coba <i>learning rate</i> pada data seimbang	72
Tabel 4.5 Hasil pengujian model dengan uji coba <i>learning rate</i> pada data seimbang	72
Tabel 4.6 Hasil pelatihan model dengan uji coba <i>epoch</i> pada data tidak seimbang	77
Tabel 4.7 Hasil pengujian model dengan uji coba <i>epoch</i> pada data tidak seimbang	78
Tabel 4.8 Hasil pelatihan model dengan uji coba <i>epoch</i> pada data seimbang	82
Tabel 4.9 Hasil pengujian model dengan uji coba <i>epoch</i> pada data seimbang	83

DAFTAR KODE PROGRAM

Kode Program 4.1 Memuat Data Citra	42
Kode Program 4.2 <i>Undersampling</i> data.....	46
Kode Program 4.3 <i>Parse</i> Data.....	47
Kode Program 4.4 <i>Resize</i> Data.....	48
Kode Program 4.5 <i>Splitting Dataset</i>	48
Kode Program 4.6 Pemanggilan model <i>transfer learning</i>	50
Kode Program 4.7 Pelatihan Model.....	52
Kode Program 4.8 Fungsi untuk membuat generator data.....	55
Kode Program 4.9 <i>History Training</i>	57
Kode Program 4.10 Pengujian model	59
Kode Program 4.11 Prediksi menggunakan model	61
Kode Program 4.12 Evaluasi <i>Confusion Matric</i> , <i>Classification Report</i> , dan <i>ROC Curve</i>	62

DAFTAR LAMPIRAN

LAMPIRAN 1 TAMPILAN PROGRAM	95
--	-----------