

## Klasifikasi Subtipe Leukemia Limfoblastik Akut (LLA) pada Citra Mikroskopis Sel Darah Menggunakan Arsitektur EfficientNet-B3 dengan Dataset Seimbang

Ni Putu Dina Agustina

Universitas Udayana, Indonesia

Email: [agustina.2208541043@student.unud.ac.id](mailto:agustina.2208541043@student.unud.ac.id)

### Abstrak:

Leukemia Limfoblastik Akut (ALL) adalah salah satu jenis kanker darah paling umum yang menyerang anak-anak dan membutuhkan diagnosis yang cepat dan akurat. Penelitian ini mengusulkan model klasifikasi untuk subtipe leukemia limfoblastik akut (ALL) berdasarkan gambar sel darah mikroskopis menggunakan arsitektur EfficientNet-B3. Dengan pendekatan pembelajaran transfer dan kumpulan data yang seimbang, model ini mencapai akurasi pengujian 97,50% dan Skor F1 rata-rata 0,97. Secara keseluruhan, nilai rata-rata makro dan rata-rata tertimbang menunjukkan hasil yang konsisten, dengan presisi dan ingatan 0,98 dan Skor F1 0,97. Ini menunjukkan bahwa model ini unggul tidak hanya dalam satu atau dua kelas tetapi menunjukkan kinerja yang seragam di semua kelas, menjadikannya alat klasifikasi yang kuat untuk aplikasi diagnosis leukemia otomatis.

**Kata kunci:** leukemia limfoblastik akut, klasifikasi citra, EfficientNet-B3, deep learning, transfer learning.

### Abstract:

*Acute Lymphoblastic Leukemia (ALL) is one of the most common types of blood cancer that affects children and requires fast and accurate diagnosis. This study proposes a classification model for subtypes of acute lymphoblastic leukemia (ALL) based on microscopic blood cell images using the EfficientNet-B3 architecture. With a transfer learning approach and a balanced dataset, the model achieves a testing accuracy of 97.50% and an average F1-Score of 0.97. Overall, the macro average and weighted average values show consistent results, with precision and recall of 0.98 and an F1-Score of 0.97. This indicates that the model excels not only in one or two classes but demonstrates uniform performance across all classes, making it a robust classification tool for automatic leukemia diagnosis applications.*

**Keywords:** acute lymphoblastic leukemia, image classification, EfficientNet-B3, deep learning, transfer learning

Corresponding: Ni Putu Dina Agustina

E-mail: [agustina.2208541043@student.unud.ac.id](mailto:agustina.2208541043@student.unud.ac.id)



## PENDAHULUAN

Leukemia merupakan jenis kanker darah yang menyerang sistem pembentukan sel darah, terutama sumsum tulang dan limfosit (S. Kumar & Qureshi, 2025). Menurut data dari Global Cancer Observatory (GCO) yang dikelola oleh International Agency for Research on Cancer (IARC) di bawah naungan WHO, pada tahun 2020 terdapat sekitar 474.519 kasus baru leukemia dan 311.594 kematian akibat penyakit ini di seluruh dunia (Ferlay et al., 2021). Kanker ini dapat menyerang berbagai kelompok usia, mulai dari anak-anak hingga dewasa. Terdapat sekitar sepertiga dari seluruh kanker yang terjadi pada anak-anak merupakan leukemia (“DOI: [Http://Dx.Doi.Org/10.33846/Sf15340](http://Dx.Doi.Org/10.33846/Sf15340) Pemberian Air Susu Ibu Eksklusif Untuk Memperkuat Sistem Imun Terhadap Risiko Kejadian Leukemia Limfoblastik Akut Pada Anak Lucy Nuryudha Ramadhani,” 2024). Hal ini menjadikan leukemia sebagai jenis kanker yang paling umum terjadi pada anak-anak (Rizki Firdaus Mulya et al., 2023) (Istiqomah et al., 2024). Khususnya tipe Leukemia Limfoblastik Akut (LLA), yang ditandai dengan proliferasi

limfoblas secara abnormal dalam darah dan sumsum tulang yang dapat mengganggu produksi sel darah normal dan memperburuk kondisi pasien secara cepat (S. Kumar & Qureshi, 2025)(Fajrina et al., 2024)(Akbar et al., 2024).

Metode diagnosis tradisional umumnya melibatkan pemeriksaan darah lengkap, biopsi sumsum tulang, dan evaluasi mikroskopis manual oleh ahli hematologi(Prajapati et al., 2024)(Istiqomah et al., 2024). Namun, metode ini memerlukan waktu, bersifat invasif, dan sangat bergantung pada keahlian subjektif manusia(Akbar et al., 2024)(Revanda et al., 2022)(Kurniadi et al., 2025). Seiring dengan berkembangnya teknologi kecerdasan buatan, *deep learning* menjadi pendekatan yang menjanjikan untuk menganalisis citra mikroskopis darah dengan akurasi tinggi dan efisiensi yang lebih baik(Istiqomah et al., 2024)(Perveen et al., 2024).

Salah satu algoritma *deep learning* paling populer untuk pengenalan pola pada citra medis adalah *Convolutional Neural Network* (CNN), yang telah digunakan dalam berbagai penelitian untuk mendeteksi dan mengklasifikasikan sel kanker darah(Rizki Firdaus Mulya et al., 2023). *Convolutional Neural Network* (CNN) memproses citra dengan menerapkan lapisan konvolusi yang bertugas mengekstraksi fitur-fitur utama seperti tepi, tekstur, dan pola. Fitur-fitur ini kemudian digabungkan melalui beberapa lapisan untuk mereduksi dimensi data, sehingga jaringan dapat mengenali pola yang lebih kompleks dan melakukan klasifikasi citra ke dalam berbagai kategori(Kurniadi et al., 2025). Beberapa studi telah membuktikan bahwa arsitektur CNN seperti VGG-16, MobileNetV2, dan InceptionV3 dapat memberikan akurasi yang sangat tinggi dalam klasifikasi leukemia(Prajapati et al., 2024)(Rizki Firdaus Mulya et al., 2023).

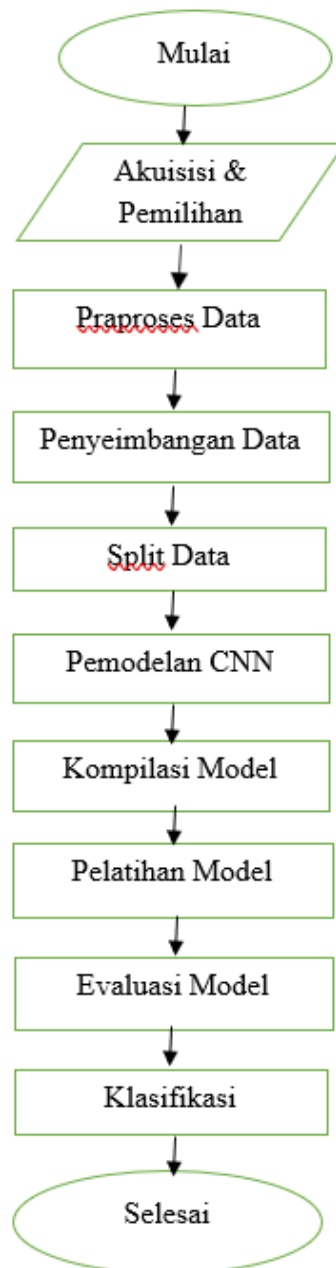
(Fajrina et al., 2024)Menunjukkan bahwa penggunaan arsitektur EfficientNet-B0 pada dataset ALL dari Kaggle mampu mencapai akurasi pengujian sebesar 98,48%. Sementara itu, penelitian (Perveen et al., 2024) mengembangkan pendekatan *deep ensemble learning* yang secara efektif mendeteksi LLA dan subtipe pada dataset yang telah di-*oversampling*. Selain itu, studi (Revanda et al., 2022) memanfaatkan teknik segmentasi berbasis Mask R-CNN untuk mendeteksi dan mengklasifikasikan subtipe ALL secara lebih rinci dari citra sel darah pasien RSUD Dr. Soetomo.

Dengan latar belakang tersebut, penelitian ini bertujuan untuk mengembangkan model klasifikasi leukemia berbasis *deep learning* dengan memanfaatkan arsitektur EfficientNet-B3. Dataset citra mikroskopis sel darah yang digunakan telah diseimbangkan menggunakan teknik *oversampling* dan *undersampling* untuk memastikan distribusi kelas yang setara. Model ini diharapkan mampu mendeteksi subtipe leukemia secara lebih akurat dan efisien, sehingga dapat mendukung proses diagnosis medis dengan cepat dan akurat.

## **METODE PENELITIAN**

### **A. Alur Penelitian**

Dalam perancangan sistem ini, disediakan gambaran umum tentang alur penelitian yang akan dikerjakan:



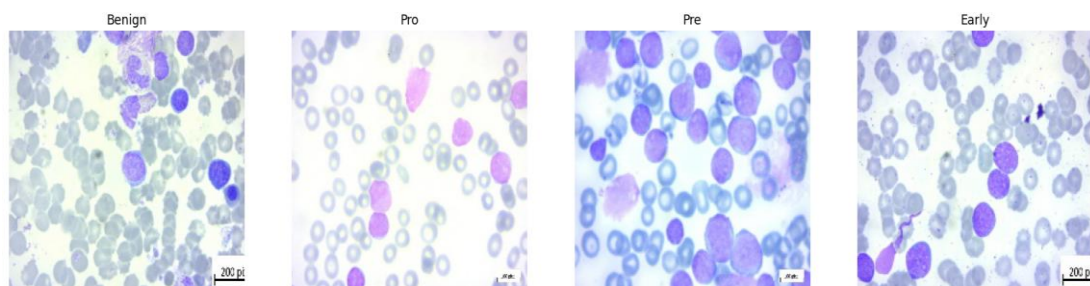
**Gambar 1. Alur Penelitian**

Klasifikasi penyakit Leukemia imfoblastik Akut (LLA) dilakukan dengan menggunakan algoritma *Convolutional Neural Network* (CNN) dengan arsitektur EfficientNet-B3 untuk mengklasifikasikan empat kelas, yaitu *Benign*, *Early*, *(Pre)* Precursor, dan *(Pro)* Progenitor. Perancangan penelitian dimulai dengan tahap akuisisi data, kemudian dilanjutkan dengan pra-pemrosesan citra, yang mencakup perubahan ukuran (*resizing*) menjadi 224×224 piksel, konversi ke RGB, serta normalisasi skala piksel sesuai standar pretrained EfficientNet-B3. Setelah itu, dilakukan penyeimbangan data antar kelas menggunakan teknik oversampling dan undersampling untuk memastikan distribusi data yang setara.

Data yang telah diproses kemudian dibagi menjadi 80% untuk *training set*, 10% untuk *validation set*, dan 10% untuk *testing set*. Tahap berikutnya adalah pembangunan dan pelatihan model melalui optimasi *hyperparameter* dan pelatihan selama sejumlah 100 *epoch*. Setelah proses pelatihan selesai, model disimpan dan dievaluasi untuk mengukur performa klasifikasinya terhadap *subtipe* leukemia berdasarkan citra mikroskopis sel darah.

## B. Akuisisi dan Pemilihan Data

Pada tahap akuisisi data, dilakukan proses pengumpulan citra yang akan digunakan untuk melatih model klasifikasi. Dalam penelitian ini, sumber data yang digunakan berasal dari dataset publik berjudul "*Acute Lymphoblastic Leukemia (ALL) image dataset*" yang dirilis oleh Mehrad Aria dan Mustafa Ghaderzadeh melalui platform Kaggle pada tahun 2021 (*Archive (1)*, 2021). Dataset ini terdiri dari citra mikroskopis sel darah yang diklasifikasikan ke dalam empat kelas: *Benign*, *Early*, *Precursor*, dan *Progenitor*. Dataset tersebut memuat total 3.256 citra mikroskopis, yang terbagi ke dalam empat kategori, yaitu 504 citra *benign*, 985 citra *early*, 963 citra *pre*, dan 804 citra *pro* (Fajrina et al., 2024). Mengingat distribusi kelas yang tidak seimbang, maka dilakukan seleksi awal dan penyesuaian jumlah data per kelas untuk mempermudah proses pelatihan model yang adil dan konsisten. Dalam penelitian ini, digunakan sebanyak 700 data dari masing-masing kelas.



Gambar 2. Kelas Dataset

## C. Praproses Data

Tahap praproses data dilakukan untuk menyiapkan citra mikroskopis agar sesuai dengan format input yang dibutuhkan oleh arsitektur model EfficientNet-B3. Proses ini mencakup beberapa langkah penting. Pertama, setiap citra diubah ukurannya menjadi  $224 \times 224$  piksel menggunakan metode *resizing*, agar sesuai dengan resolusi input default dari EfficientNet-B3. Selain itu, konversi format warna ke RGB dilakukan untuk memastikan keseragaman saluran warna. Langkah ini penting karena beberapa gambar pada dataset memiliki saluran warna berbeda atau tidak lengkap. Setelah itu, dilakukan normalisasi piksel menggunakan fungsi *preprocess\_input* dari pustaka *TensorFlow*, yang bertujuan menyelaraskan distribusi nilai piksel dengan standar data *pretrained ImageNet*. Normalisasi ini membantu model memahami pola visual secara lebih efektif dan mempercepat proses pelatihan.

Teknik praproses seperti ini umum digunakan dalam pemodelan berbasis citra dan terbukti efektif dalam meningkatkan akurasi model klasifikasi, seperti yang juga dijelaskan

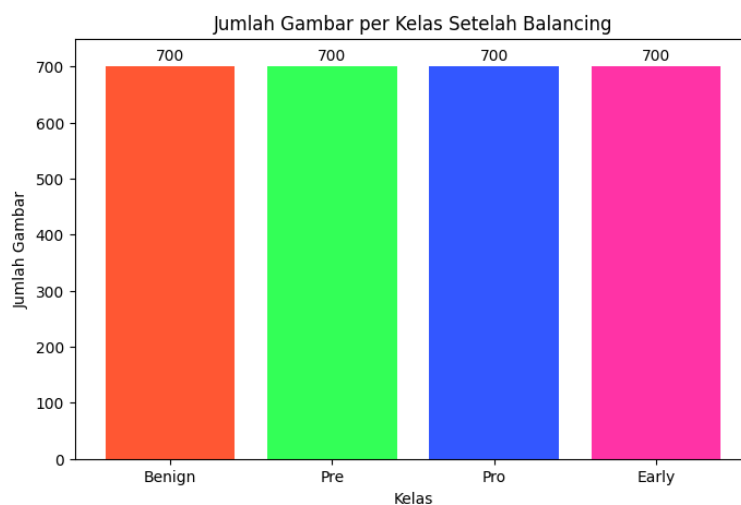
pada penelitian Gori et al. (2024), yang menunjukkan bahwa proses integrasi, transformasi, dan pembersihan data memiliki pengaruh signifikan terhadap kinerja model machine learning, khususnya dalam mengurangi *noise* dan meningkatkan konsistensi data input (Gori et al., 2024).

Dengan demikian, tahapan praproses pada penelitian ini berfungsi untuk menyederhanakan struktur data citra dan memastikan bahwa setiap input memiliki format yang seragam, sehingga proses ekstraksi fitur oleh model dapat berjalan lebih optimal.

#### D. Penyeimbangan Data

Penyeimbangan data dilakukan untuk mengatasi ketidakseimbangan distribusi jumlah citra pada masing-masing kelas dalam dataset. Pada kondisi awal, dataset terdiri dari 504 citra *benign*, 985 citra *early*, 963 citra *pre*, dan 804 citra *pro*. Ketimpangan ini berisiko menimbulkan bias model terhadap kelas mayoritas, sehingga dapat menurunkan akurasi prediksi terhadap kelas minoritas.

Untuk mengatasi masalah tersebut, dilakukan teknik kombinasi *oversampling* dan *undersampling* secara terprogram, dengan tujuan menyamakan jumlah data di keempat kelas menjadi masing-masing 700 citra. *Oversampling* dilakukan dengan cara menduplikasi citra pada kelas minoritas, sementara *undersampling* dilakukan dengan memilih sebagian citra dari kelas mayoritas secara acak. Penyeimbangan ini menghasilkan total dataset sebanyak 2.800 citra.



Gambar 3. Dataset Seimbang

Teknik ini dipilih karena terbukti mampu memperbaiki distribusi data dan meningkatkan performa model klasifikasi pada data yang semula tidak seimbang. Menurut Erlangga et al. (2024), kombinasi metode *oversampling* dan *undersampling* dapat meningkatkan kinerja model klasifikasi dengan mengurangi dominasi kelas mayoritas dan mengatasi masalah *overlapping* antar kelas (Erlangga et al., 2024).

Dengan demikian, penyeimbangan dataset merupakan langkah penting untuk memastikan bahwa proses pelatihan model tidak bias terhadap salah satu kelas, serta mendukung evaluasi performa yang lebih akurat dan representatif.

#### E. Split Data

Setelah data selesai melalui tahap praproses dan penyeimbangan, dataset dibagi menjadi dua bagian utama, yaitu 80% sebagai data latih dan 20% sebagai data uji. Pembagian ini bertujuan untuk memastikan bahwa model dilatih dan dievaluasi menggunakan data yang berbeda, sehingga dapat mengukur kemampuan generalisasi model secara objektif dan menghindari overfitting. Strategi ini merupakan pendekatan umum dalam pengembangan sistem kecerdasan buatan di bidang kesehatan, sebagaimana diterapkan dalam penelitian oleh Kumar dan Qureshi (2025), di mana pembagian dataset menjadi data pelatihan dan pengujian digunakan untuk meningkatkan validitas hasil diagnosis sistem berbasis AI dalam klasifikasi kasus leukemia (S. Kumar & Qureshi, 2025).

### F. Model Deep Learning

Model deep learning yang digunakan dalam penelitian ini adalah arsitektur EfficientNet-B3, yang tergolong dalam kelompok *Convolutional Neural Network* (CNN) berbasis transfer learning. CNN merupakan algoritma populer dalam pengolahan citra karena kemampuannya dalam mengekstraksi fitur visual seperti tepi, tekstur, hingga pola kompleks secara hierarkis melalui lapisan konvolusinya (Adiwijaya et al., 2025). Arsitektur CNN telah banyak diterapkan dalam klasifikasi penyakit berbasis citra medis, termasuk untuk deteksi leukemia dari citra darah tepi maupun sumsum tulang (Claro et al., 2020). Misalnya, penelitian oleh Kumar et al. (2020) menunjukkan bahwa CNN mampu secara akurat mengenali kanker darah putih dari citra mikroskopis sumsum tulang (D. Kumar et al., 2020). Dalam konteks yang lebih spesifik, Shoumi et al. (2023) mengembangkan model Mask R-CNN untuk deteksi sel plasma pada kasus multiple myeloma, menunjukkan bahwa pendekatan CNN dan turunannya sangat relevan untuk permasalahan klasifikasi berbasis sel mikroskopis (Milyun Ni'ma Shoumi et al., 2023).

### G. Arsitektur EfficientNetB3

EfficientNet merupakan arsitektur *Convolutional Neural Network* (CNN) yang dikembangkan oleh Google AI dengan pendekatan *compound scaling*, yaitu metode penskalaan terkoordinasi terhadap tiga aspek utama dalam jaringan saraf, yakni kedalaman (*depth*), lebar (*width*), dan resolusi *input*. Pendekatan ini memungkinkan pengembangan model yang lebih akurat sekaligus efisien dalam jumlah parameter dan kebutuhan komputasi.



#### Gambar 4. Representasi Arsitektur EfficientNetB3

Dalam penelitian ini, digunakan varian EfficientNet-B3 yang telah dipra-latih menggunakan dataset *ImageNet* sebagai *feature extractor*. Dibandingkan dengan varian EfficientNet-B0 yang lebih ringan, EfficientNet-B3 memiliki kapasitas lebih tinggi dalam mengekstraksi fitur visual kompleks dari citra mikroskopis, sehingga lebih sesuai untuk tugas klasifikasi citra medis seperti leukemia. Studi sebelumnya oleh Anggiratih et al. (2021) menunjukkan keberhasilan arsitektur ini dalam klasifikasi penyakit tanaman berbasis citra, sedangkan Andre et al. (2021) membuktikan efektivitasnya dalam klasifikasi tumor otak menggunakan pendekatan transfer learning (Anggiratih et al., 2021)(Andre et al., 2021).

Model EfficientNet-B3 dalam penelitian ini dimodifikasi dengan penambahan beberapa lapisan pada bagian atas model. Lapisan *GlobalAveragePooling2D* digunakan untuk mereduksi dimensi spasial hasil ekstraksi fitur tanpa kehilangan informasi penting. Selanjutnya, ditambahkan lapisan *dropout* untuk mengurangi risiko *overfitting* selama pelatihan. Sebagai penutup, digunakan satu lapisan *dense* dengan aktivasi *softmax* yang berfungsi mengklasifikasikan input ke dalam empat kelas yang ditentukan, yaitu *Benign*, *Early*, *Precursor*, dan *Progenitor*.

Fungsi *loss* yang digunakan adalah *sparse categorical crossentropy*, karena label kelas yang digunakan berbentuk numerik (*integer*), bukan *one-hot encoding*. Selain itu, proses pelatihan model menggunakan algoritma Adam sebagai *optimizer*, yang dikenal karena kemampuannya melakukan pembaruan bobot secara efisien dan stabil pada jaringan dalam. Penggunaan konfigurasi ini bertujuan memaksimalkan akurasi model sambil mempertahankan kecepatan konvergensi dan efisiensi proses pelatihan.

#### H. Hyperparameter

*Hyperparameter* merupakan parameter eksternal yang ditetapkan sebelum proses pelatihan model dimulai. Dalam penelitian ini, beberapa *hyperparameter* yang digunakan meliputi jenis *optimizer*, nilai *learning rate*, jumlah *epoch* pelatihan, serta ukuran *batch* yang digunakan selama proses training berlangsung.

##### 1. *Optimizer*

Pada penelitian ini digunakan algoritma optimisasi adam sebagai *optimizer*, karena dikenal adaptif dan efisien dalam mempercepat konvergensi model selama pelatihan. Adam menggabungkan keunggulan dari metode momentum dan *RMSProp*, sehingga dapat menyesuaikan laju pembelajaran setiap parameter secara dinamis. Dalam konteks klasifikasi citra sel darah, penggunaan Adam telah terbukti memberikan hasil yang stabil dan akurat, sebagaimana dibuktikan oleh Kurniadi et al. (2025) dalam studi perbandingan berbagai jenis optimizer pada model CNN untuk klasifikasi sel darah putih, di mana adam menunjukkan kinerja unggul dalam hal akurasi dan kestabilan pelatihan (Kurniadi et al., 2025).

##### 2. *Learning rate*

*Learning rate* merupakan nilai yang menentukan seberapa besar langkah pembaruan bobot dalam setiap iterasi pelatihan. Nilai *learning rate* yang terlalu besar dapat menyebabkan model gagal konvergen, sedangkan nilai yang terlalu kecil

memperlambat proses pelatihan. Dalam penelitian ini digunakan nilai  $1e-4$ , yang dipilih berdasarkan eksperimen awal untuk menjaga kestabilan pelatihan.

### 3. *Epoch*

*Epoch* adalah jumlah siklus penuh di mana seluruh data latih diproses oleh model selama pelatihan. Pada penelitian ini, jumlah *epoch* maksimum ditetapkan sebanyak 100, namun dengan mekanisme *early stopping* yang memungkinkan pelatihan dihentikan lebih awal jika tidak terjadi peningkatan akurasi validasi dalam beberapa iterasi. Pendekatan ini bertujuan untuk mencegah *overfitting* dan mengoptimalkan waktu pelatihan.

### 4. *Batch size*

*Batch size* menentukan jumlah data yang diproses sekaligus dalam satu iterasi pelatihan. Dalam penelitian ini digunakan *batch size* sebesar **32**, yang dianggap seimbang antara efisiensi komputasi dan kestabilan pembaruan bobot. Istiqomah et al. (2024) juga menggunakan *batch size* serupa dalam klasifikasi kanker darah berbasis CNN dan melaporkan performa pelatihan yang optimal (Istiqomah et al., 2024).

## I. Skenario Pengujian

Skenario pengujian *hyperparameter* merupakan proses penentuan konfigurasi terbaik untuk komponen penting dalam pelatihan model *deep learning*. Meskipun dalam penelitian ini tidak dilakukan pengujian kombinasi *hyperparameter* secara eksplisit seperti *grid search*, pemilihan konfigurasi dilakukan melalui proses eksplorasi bertahap berdasarkan kestabilan model dan hasil evaluasi pada data uji. Tiga komponen utama yang menjadi fokus dalam penentuan *hyperparameter* adalah sebagai berikut:

### 1. Pengujian *Optimizer*

Model dikembangkan menggunakan *optimizer* Adam, yang dipilih berdasarkan hasil pelatihan awal karena memberikan akurasi yang stabil dan konvergensi cepat. Tidak dilakukan perbandingan dengan *optimizer* lain secara eksplisit, tetapi pemilihan Adam didasarkan pada studi terdahulu dan performa yang konsisten selama pelatihan.

### 2. Pengujian *Learning rate*

Nilai *learning rate* yang digunakan dalam penelitian ini adalah  $1e-4$  atau 0.0001, yang dipilih berdasarkan uji coba awal untuk menghindari ketidakstabilan pelatihan. Nilai ini terbukti menjaga konsistensi penurunan *loss* dan peningkatan akurasi tanpa lonjakan ekstrem.

### 3. Pengujian *Epoch*

Jumlah *epoch* maksimum ditetapkan sebanyak 100, dengan diterapkannya teknik *early stopping* untuk menghentikan pelatihan secara otomatis jika akurasi validasi tidak mengalami peningkatan. Strategi ini dipilih untuk menghindari *overfitting* dan mempercepat waktu pelatihan secara efisien.

Tujuan dari skenario ini adalah menemukan kombinasi *hyperparameter* yang mampu menghasilkan model klasifikasi leukemia dengan akurasi tinggi dan nilai *loss* yang rendah. Strategi serupa juga diterapkan oleh Perveen et al. (2024) dalam penelitian klasifikasi leukemia menggunakan pendekatan deep ensemble learning, yang

menunjukkan pentingnya pemilihan konfigurasi pelatihan secara tepat dalam menghasilkan performa model terbaik(Perveen et al., 2024).

### J. Confusion Matrix

*Confusion matrix* merupakan metode evaluasi yang digunakan untuk menggambarkan kinerja model klasifikasi dalam membedakan masing-masing kelas. Matriks ini menampilkan jumlah prediksi yang benar maupun salah untuk setiap kelas dalam bentuk tabel dua dimensi, yang sangat membantu dalam mengevaluasi kekuatan model secara rinci. Dalam penelitian ini, *confusion matrix* digunakan untuk mengevaluasi model EfficientNet-B3 terhadap empat kelas leukemia, yaitu *Benign*, *Early*, *Precursor*, dan *Progenitor*. Pendekatan ini juga digunakan oleh Prajapati et al. (2024) dalam klasifikasi kanker darah menggunakan CNN, di mana *confusion matrix* memberikan gambaran menyeluruh terhadap distribusi prediksi model(Prajapati et al., 2024).

**Tabel 1. Confusion Matrix**

	Aktual Positif	Aktual Negatif
Prediksi Positif	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
Prediksi Negatif	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

Beberapa parameter digunakan untuk mengukur kinerja sistem yang akan dievaluasi dalam studi ini, yaitu:

#### 1. Accuracy

*Accuracy* merupakan metrik evaluasi yang paling dasar dalam klasifikasi, yang mengukur proporsi prediksi yang benar dibandingkan dengan keseluruhan jumlah data uji. Dalam penelitian ini, *accuracy* digunakan untuk menilai seberapa baik model EfficientNet-B3 mengklasifikasikan keempat kelas leukemia secara keseluruhan. Meskipun berguna, akurasi tidak selalu mencerminkan performa model secara menyeluruh, terutama saat data tidak seimbang. Namun demikian, karena dataset telah diseimbangkan sebelumnya, *accuracy* tetap menjadi indikator utama performa model. Pendekatan serupa digunakan oleh Claro et al. (2020), yang memanfaatkan *accuracy* untuk mengevaluasi kinerja CNN dalam diagnosis leukemia akut(Claro et al., 2020). Adapun rumus untuk menghitung *accuracy* ditunjukkan oleh persamaan (1) berikut.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

#### 2. Precision

*Precision* mengukur proporsi prediksi positif yang benar, yaitu perbandingan antara *true positive* dan jumlah seluruh prediksi positif ( $TP + FP$ ), sehingga mencerminkan ketepatan model dalam mengidentifikasi kasus positif. Adapun rumus untuk menghitung *precision* ditunjukkan oleh persamaan (2) berikut.

$$precision = \frac{TP}{TP + FP} \tag{2}$$

### 3. Recall

*Recall* adalah metrik evaluasi yang digunakan untuk mengukur kemampuan model dalam menemukan seluruh kasus positif yang benar. Nilainya dihitung sebagai rasio antara jumlah prediksi positif yang benar (*true positive*) terhadap total seluruh kasus positif yang sebenarnya (*true positive + false negative*). Adapun rumus untuk menghitung *recall* ditunjukkan oleh persamaan (3) berikut.

$$\begin{aligned} & recall \\ &= \frac{TP}{TP + FN} \end{aligned} \quad (3)$$

### 4. F1-Score

*F1-Score* adalah metrik evaluasi yang menggabungkan *precision* dan *recall* ke dalam satu nilai harmonis, terutama ketika distribusi kelas tidak seimbang atau ketika penting untuk mempertimbangkan keduanya secara bersamaan. Adapun rumus untuk menghitung *F1-score* ditunjukkan oleh persamaan (4) berikut.

$$\begin{aligned} & F1 - score \\ &= 2 \times \frac{precision \times recall}{precision + recall} \end{aligned} \quad (4)$$

### 5. Loss

Parameter *loss* merepresentasikan tingkat kesalahan sistem dalam melakukan prediksi terhadap klasifikasi suatu data masukan (*input*).

## HASIL DAN PEMBAHASAN

Hasil dari penerapan arsitektur EfficientNetB3 pada dataset Leukemia Limfoblastik Akut (LLA) (Archive (1), 2021).

### A. Analisis Hasil Pengujian *Hyperparameter* pada model *Deep Learning*

**Tabel 2. Hasil Pengujian**

<i>Epoch</i>	<i>Data Train</i>		<i>Data Validation</i>	
	<i>Loss</i>	<i>Acc</i>	<i>Val Los</i>	<i>Val Acc</i>
1	0.2584	0.9332	0.2017	0.9643
10	0.2544	0.9292	0.1732	0.9750
20	0.2134	0.9437	0.1508	0.9786
30	0.1918	0.9485	0.1319	0.9821
40	0.1859	0.9537	0.1223	0.9821
50	0.1663	0.9567	0.1055	0.9821
60	0.1655	0.9587	0.0949	0.9821
70	0.1510	0.9600	0.0932	0.9821
80	0.1429	0.9624	0.0858	0.9821
90	0.1340	0.9680	0.0818	0.9857
100	0.1412	0.9619	0.0773	0.9857

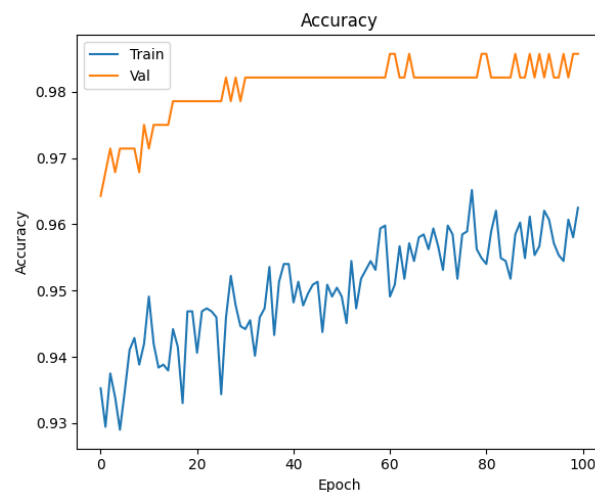
Tabel 2 menunjukkan hasil pengujian model *deep learning* selama 100 *epoch*. Dari tabel tersebut, terlihat bahwa nilai *loss* pada data *train* dan *validation* cenderung mengalami penurunan yang konsisten seiring bertambahnya jumlah *epoch*. Pada *epoch* ke-1, *train loss* tercatat sebesar 0.2584, dan secara bertahap menurun menjadi 0.1340 pada *epoch* ke-90, sebelum sedikit meningkat menjadi 0.1412 di *epoch* ke-100. Sementara itu,

*validation loss* mengalami penurunan dari 0.2017 pada *epoch* ke-1 menjadi 0.0773 di *epoch* ke-100.

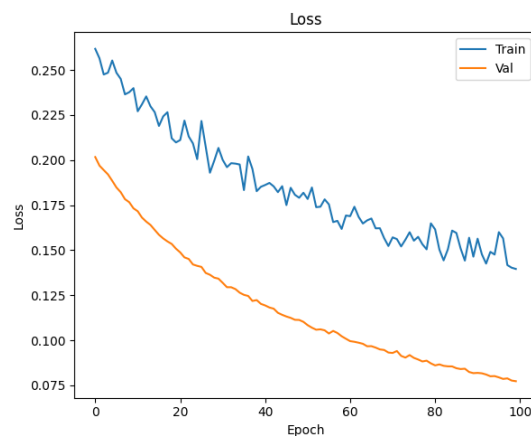
Penurunan nilai *loss* ini menunjukkan bahwa model mampu mempelajari pola dalam data secara efektif. Selain itu, akurasi pada data *train* meningkat dari 93.32% pada *epoch* ke-1 menjadi 96.80% pada *epoch* ke-90, sebelum sedikit menurun menjadi 96.19% pada *epoch* ke-100. Akurasi *validation* pun meningkat dari 96.43% menjadi 98.57%, dan relatif stabil setelah *epoch* ke-40.

Stabilitas *validation accuracy* sejak *epoch* ke-40 hingga 100 menunjukkan bahwa model tidak mengalami *overfitting* yang berarti, meskipun terus dilatih lebih lama. Hal ini mengindikasikan bahwa model mampu melakukan *generalization* dengan baik terhadap data yang tidak terlihat (*unseen data*).

Peningkatan performa ini tidak lepas dari pemanfaatan metode *transfer learning* menggunakan arsitektur EfficientNet-B3, yang terbukti mampu mempercepat konvergensi model serta menghasilkan akurasi tinggi dengan nilai *loss* yang rendah. Hasil validasi akhir yang mencapai akurasi 98.57% menunjukkan bahwa kombinasi antara kualitas dataset dan arsitektur model sangat mendukung pencapaian performa yang optimal. Visualisasi hasil pelatihan dan validasi dapat dilihat pada gambar di bawah ini.



**Gambar 5. Training dan validation loss**



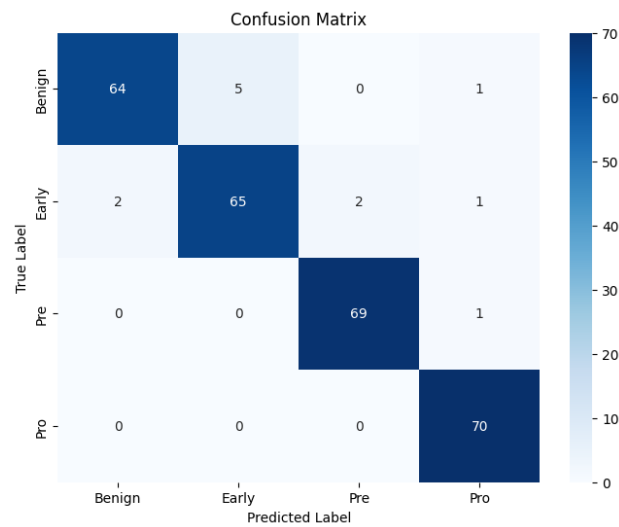
**Gambar 6. Training dan validation accuracy**

Pada Gambar 5, terlihat bahwa akurasi model terhadap data pelatihan (*train*) dan data validasi (*val*) terus meningkat seiring bertambahnya jumlah *epoch*. Kurva akurasi validasi bahkan cenderung lebih tinggi dan lebih stabil dibandingkan kurva pelatihan, yang menunjukkan bahwa model tidak hanya menunjukkan kinerja tinggi terhadap data pelatihan, tetapi juga sangat akurat dalam mengklasifikasikan data yang belum pernah dilihat sebelumnya. Fenomena ini menandakan bahwa model tidak mengalami *overfitting* dan mampu melakukan generalisasi dengan sangat baik. Kurva validasi yang lebih tinggi dari pelatihan pada sebagian besar *epoch* juga mengindikasikan bahwa model benar-benar berhasil mempelajari pola-pola relevan dari fitur data, bukan sekadar menghafalnya.

Sementara itu, pada Gambar 6, kurva *loss* untuk data pelatihan dan validasi menunjukkan tren penurunan yang konsisten dan paralel. *Validation loss* secara konsisten berada di bawah *train loss*, yang semakin menguatkan bahwa model memiliki kemampuan generalisasi yang sangat baik. Tidak tampak fluktuasi tajam atau peningkatan *validation loss* di akhir-akhir *epoch*, sehingga bisa disimpulkan bahwa proses pelatihan berlangsung stabil dan model tidak mengalami *overfitting*.

Secara keseluruhan, visualisasi ini menunjukkan bahwa model yang dibangun menggunakan arsitektur EfficientNet-B3 berhasil melakukan proses pembelajaran yang efisien dan akurat, dengan kinerja tinggi pada kedua subset data.

## **B. Confusion Matrix**



**Gambar 7. Confusion Matrix**

Berdasarkan confusion matrix pada Gambar 7, model menunjukkan performa klasifikasi yang sangat baik terhadap keempat kelas: *Benign*, *Early*, *Pre*, dan *Pro*. Kelas *Pro* berhasil diklasifikasikan secara sempurna tanpa kesalahan (70/70), menunjukkan bahwa model memiliki sensitivitas dan spesifisitas yang sangat tinggi terhadap kategori ini. Sementara itu, kelas *pre* hanya mengalami satu kesalahan klasifikasi, dengan 69 dari 70 sampel diklasifikasikan dengan benar, mencerminkan kemampuan model yang sangat baik dalam membedakan tahap pra-kanker.

Untuk kelas *Benign*, sebanyak 64 dari 70 sampel berhasil diklasifikasikan dengan benar, dengan sebagian kecil (5 sampel) salah diklasifikasikan sebagai *Early* dan 1 sampel sebagai *Pro*. Demikian pula, pada kelas *Early*, terdapat 65 prediksi yang tepat dari 70 sampel, dengan 2 sampel salah diklasifikasikan sebagai *Benign*, 2 sebagai *Pre*, dan 1 sebagai *Pro*.

Secara keseluruhan, distribusi kesalahan klasifikasi tergolong sangat kecil dan tidak bersifat sistematis, yang mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik dalam membedakan fitur visual antar kelas. Hasil ini memperkuat validitas arsitektur yang digunakan, dalam hal ini EfficientNet-B3, dalam menangani tugas klasifikasi multi-kelas pada citra medis.

### C. Evaluasi Model

**Tabel 3. Hasil Evaluasi Model**

<i>Score</i>	<i>Loss</i>	<i>Accuracy</i>
<i>Train</i>	0.1396	0.9625
<i>Validation</i>	0.0773	0.9857
<i>Test</i>	0.1093	0.9750

Tabel 3 menyajikan hasil evaluasi performa model pada data pelatihan, validasi, dan pengujian. Model menunjukkan kinerja yang sangat baik pada data pelatihan, dengan nilai *loss* sebesar 0.1396 dan akurasi mencapai 96.25%, yang mengindikasikan bahwa

model mampu belajar secara efektif dari data yang diberikan. Pada data validasi, performa model meningkat, ditandai dengan penurunan nilai *loss* menjadi 0.0773 dan peningkatan akurasi hingga 98.57%. Hal ini menunjukkan kemampuan generalisasi model yang sangat baik terhadap data yang belum pernah dilihat sebelumnya selama pelatihan. Sementara itu, evaluasi pada data uji menghasilkan nilai *loss* sebesar 0.1093 dan akurasi 97.50%. Meskipun sedikit lebih rendah dibandingkan dengan akurasi validasi, performa ini tetap menunjukkan bahwa model memiliki generalisasi yang kuat dan stabil saat diterapkan pada data baru. Secara keseluruhan, hasil evaluasi ini mengindikasikan bahwa model tidak mengalami overfitting dan mampu mempertahankan kinerja tinggi pada seluruh subset data.

**Tabel 4. Laporan Klasifikasi**

	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
<i>Benign</i>	0.97	0.96	0.96	70
<i>Pre</i>	1.00	0.99	0.99	70
<i>Pro</i>	0.96	1.00	0.98	70
<i>Early</i>	0.97	0.96	0.96	70
<i>Accuracy</i>	-	-	0.97	280
<i>Macro avg</i>	0.98	0.98	0.97	280
<i>Weighted avg</i>	0.98	0.97	0.97	280

Berdasarkan hasil evaluasi performa klasifikasi pada Tabel 4, model berhasil mencapai akurasi keseluruhan sebesar 97% dalam membedakan empat kelas leukemia, yaitu *Benign*, *Pre*, *Pro*, dan *Early*. Nilai *precision* tertinggi diperoleh oleh kelas *pre* dan *precision* terendah diperoleh oleh kelas *pro*. Parameter *recall* tertinggi terjadi pada kelas *pro* dengan nilai 1.00 sedangkan nilai parameter *recall* terendah terjadi pada kelas *benign* dan *early* dengan nilai yang sama yaitu 0.96. Kemudian untuk parameter *f1-score*, nilai tertinggi diperoleh oleh kelas *pro* dan nilai terendah terjadi pada kelas *benign*. Dengan nilai *support* yang sama pada masing-masing kelas yaitu 70. Hal ini berarti ada 70 data uji yang digunakan dalam setiap kelas. Nilai *support* yang seimbang antar kelas membantu model belajar secara adil terhadap seluruh kelas, sementara nilai *support* yang tidak seimbang dapat menyebabkan model bias terhadap kelas dengan jumlah data lebih besar. Oleh karena itu, penting untuk memperhatikan distribusi *support* saat mengevaluasi performa model klasifikasi.

## KESIMPULAN

Berdasarkan penelitian dan evaluasi model yang telah dilakukan, maka dapat disimpulkan bahwa penelitian ini berhasil membangun model klasifikasi leukemia limfoblastik akut (LLA) dari citra mikroskopis sel darah dengan memanfaatkan arsitektur EfficientNet-B3 serta strategi penyeimbangan dataset antar kelas. Melalui tahapan praproses yang tepat dan konfigurasi hyperparameter yang optimal, model yang dikembangkan mampu mencapai akurasi pelatihan sebesar 96.25%, akurasi validasi sebesar 98.57%, dan akurasi pengujian

sebesar 97.50% dan F1-Score rata-rata 0.97. Hasil ini menegaskan bahwa pendekatan *deep learning* dengan EfficientNet-B3 merupakan solusi yang efektif dan efisien dalam klasifikasi subtipe LLA berbasis citra. Untuk pengembangan selanjutnya, disarankan agar model diuji pada dataset yang lebih bervariasi, serta diperkaya dengan teknik augmentasi data atau pendekatan *ensemble learning* guna meningkatkan kemampuan generalisasi dan adaptasi model terhadap data dunia nyata.

## DAFTAR PUSTAKA

- Adiwijaya, N. O., Ardiansyah, S., & Swasono, D. I. (2025). Detecting acute lymphoblastic leukemia in blood smear images using CNN and SVM. *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control Journal*, 10(1), 29–36. <https://doi.org/https://doi.org/10.22219/kinetik.v10i1.2027>
- Akbar, M., Prasetyaningrum, P. T., Setyaningsih, P. W., Ahsan, M., & Budianto, A. E. (2024). Deteksi Leukemia Limfoblastik Akut menggunakan Convolutional Neural Network. *Jurnal Teknologi Sistem Informasi Dan Aplikasi*, 7(1), 292–297. <https://doi.org/10.32493/jtsi.v7i1.34168>
- Andre, R., Wahyu, B., & Purbaningtyas, R. (2021). Klasifikasi Tumor Otak Menggunakan Convolutional Neural Network Dengan Arsitektur Efficientnet-B3. *Jurnal IT*, 11(3), 55–59.
- Anggiratih, E., Siswanti, S., Octaviani, S. K., & Sari, A. (2021). Klasifikasi Penyakit Tanaman Padi Menggunakan Model Deep Learning Efficientnet B3 dengan Transfer Learning. *Jurnal Ilmiah SINUS*, 19(1), 75. <https://doi.org/10.30646/sinus.v19i1.526>
- Archive (1). (2021).
- Claro, M., Vogado, L., Veras, R., Santana, A., Tavares, J., Santos, J., & MacHado, V. (2020). Convolution Neural Network Models for Acute Leukemia Diagnosis. *International Conference on Systems, Signals, and Image Processing, 2020-July*, 63–68. <https://doi.org/10.1109/IWSSIP48289.2020.9145406>
- DOI: <http://dx.doi.org/10.33846/sf15340> Pemberian Air Susu Ibu Eksklusif untuk Memperkuat Sistem Imun terhadap Risiko Kejadian Leukemia Limfoblastik Akut pada Anak Lucy Nuryudha Ramadhani. (2024). *Jurnal Penelitian Kesehatan Suara Forikes*, 15(September), 538–542. <https://doi.org/http://dx.doi.org/10.33846/sf15340>
- Erlangga, A. A. G. W. S., Gunadi, I. G. A., & Sunarya, I. M. G. (2024). Kombinasi Oversampling dan Undersampling dalam Menangani Class Imbalanced dan Overlapping pada Klasifikasi Data Bank Marketing. *Jurnal RESISTOR (Rekayasa Sistem Komputer)*, 7(1), 32–42. <https://doi.org/10.31598/jurnalresistor.v7i1.1515>
- Fajrina, A. N., Pradana, Z. H., Purnama, S. I., & Romadhona, S. (2024). Penerapan Arsitektur EfficientNet-B0 Pada Klasifikasi Leukimia Tipe Acute Lymphoblastik Leukimia. *Jurnal Riset Rekayasa Elektro*, 6(1), 59–68. <https://doi.org/10.30595/jrre.v6i1.22090>
- Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International Journal of Cancer*, 149(4), 778–789. <https://doi.org/10.1002/ijc.33588>
- Gori, T., Sunyoto, A., & Al Fatta, H. (2024). Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(1), 215–224. <https://doi.org/10.25126/jtiik.20241118074>
- Istiqomah, H., Purwono, P., & Ardianto, R. (2024). Prediksi Kanker Darah Menggunakan Metode Convolutional Neural Network. *Jurnal Ilmu Komputer Dan Informatika*, 4(1), 51–60. <https://doi.org/10.54082/jiki.156>

- Kumar, D., Jain, N., Khurana, A., Mittal, S., Satapathy, S. C., Senkerik, R., & Hemanth, J. D. (2020). Automatic Detection of White Blood Cancer from Bone Marrow Microscopic Images Using Convolutional Neural Networks. *IEEE Access*, 8, 142521–142531. <https://doi.org/10.1109/ACCESS.2020.3012292>
- Kumar, S., & Qureshi, H. A. (2025). Transforming Pediatric Leukemia Care : The Role of Artificial Intelligence in Diagnosing , Treating , and Optimizing Outcomes in Acute Lymphoblastic Leukemia. *Journal of Development and Social Sciences*, 6(1), 348–363. [https://doi.org/https://doi.org/10.47205/jdss.2025\(6-1\)31](https://doi.org/https://doi.org/10.47205/jdss.2025(6-1)31)
- Kurniadi, D., Shidiq, R. M., & Mulyani, A. (2025). *Perbandingan Penggunaan Optimizer dalam Klasifikasi Sel Darah Putih Menggunakan Convolutional Neural Network*. 14(1), 77–86. <https://doi.org/10.22146/jnteti.v14i1.17162>
- Milyun Ni'ma Shoumi, Radian Malek Rayrendra, Dwi Puspitasari, & Pramana Yoga Saputra. (2023). Plasma Cell Detection in Multiple Myeloma Cases Using Mask Region Based Convolutional Neural Network Method (Mask R-CNN). *Jurnal Nasional Pendidikan Teknik Informatika (JANAPATI)*, 12(1), 156–163. <https://doi.org/10.23887/janapati.v12i1.53119>
- Perveen, S., Alourani, A., Shahbaz, M., Ashraf, M. U., & Hamid, I. (2024). A Framework for Early Detection of Acute Lymphoblastic Leukemia and Its Subtypes From Peripheral Blood Smear Images Using Deep Ensemble Learning Technique. *IEEE Access*, 12, 29252–29268. <https://doi.org/10.1109/ACCESS.2024.3368031>
- Prajapati, M., Baliarsingh, S. K., Hota, J., Dev, P. P., & Das, S. (2024). Deteksi dan Klasifikasi Citra Kanker Darah Menggunakan Metode Convolutional Neural Network (CNN). *Lecture Notes in Electrical Engineering*, 7(4), 159–165. [https://doi.org/10.1007/978-981-99-7077-3\\_16](https://doi.org/10.1007/978-981-99-7077-3_16)
- Revanda, A. R., Faticah, C., & Suciati, N. (2022). Classification of Acute Lymphoblastic Leukemia on White Blood Cell Microscopy Images Based on Instance Segmentation Using Mask R-CNN. *International Journal of Intelligent Engineering and Systems*, 15(5), 625–637. <https://doi.org/10.22266/ijies2022.1031.54>
- Rizki Firdaus Mulya, Ema Utami, & Dhani Ariatmanto. (2023). Classification of Acute Lymphoblastic Leukemia based on White Blood Cell Images using InceptionV3 Model. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 7(4), 947–952. <https://doi.org/10.29207/resti.v7i4.5182>