

**KLASIFIKASI *ACUTE LYMPHOBLASTIC LEUKEMIA* (ALL)
MENGUNAKAN *SQUEEZENET* BERDASARKAN
CITRA SEL DARAH**

SKRIPSI

Oleh :
ZAINAL ABIDIN
NIM. 220605110025



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

**KLASIFIKASI *ACUTE LYMPHOBLASTIC LEUKEMIA* (ALL)
MENGUNAKAN *SQUEEZENET* BERDASARKAN
CITRA SEL DARAH**

SKRIPSI

Diajukan kepada:
Universitas Islam Negeri Maulana Malik Ibrahim Malang
Untuk memenuhi Salah Satu Persyaratan dalam
Memperoleh Gelar Sarjana Komputer (S.Kom)

Oleh :
ZAINAL ABIDIN
NIM. 220605110025

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM
MALANG
2026**

HALAMAN PERSETUJUAN

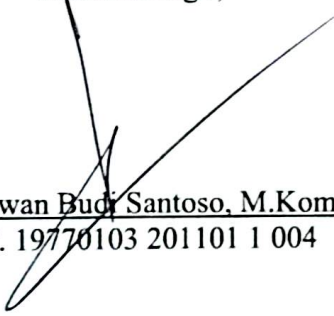
KLASIFIKASI ACUTE LYMPHOBLASTIC LEUKEMIA (ALL) MENGUNAKAN SQUEEZENET BERDASARKAN CITRA SEL DARAH

SKRIPSI

Oleh :
ZAINAL ABIDIN
NIM. 220605110025

Telah Diperiksa dan Disetujui untuk Diuji:
Tanggal: 15 Juni 2026

Pembimbing I,


Dr. Irwan Budi Santoso, M.Kom
NIP. 19770103 201101 1 004

Pembimbing II,


Khadijah Fahmi Hayati Holle, M.Kom
NIP. 19900626 202203 2 002

Mengetahui,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang



Supriyanto, M. Kom
NIP. 19890110 201903 1 012

HALAMAN PENGESAHAN

KLASIFIKASI ACUTE LYMPHOBLASTIC LEUKEMIA (ALL) MENGUNAKAN SQUEEZENET BERDASARKAN CITRA SEL DARAH

SKRIPSI

Oleh :

ZAINAL ABIDIN
NIM. 220605110025

Telah Dipertahankan di Depan Dewan Penguji Skripsi
dan Dinyatakan Diterima Sebagai Salah Satu Persyaratan
Untuk Memperoleh Gelar Sarjana Komputer (S.Kom)
Tanggal: ~~22~~ Juni 2026

Susunan Dewan Penguji

Ketua Penguji : Prof. Dr. Suhartono S.Si M.Kom
NIP. 196805192003121001

Anggota Penguji I : Okta Qomaruddin Aziz, M.Kom
NIP. 199110192019031013

Anggota Penguji II : Dr. Irwan Budi Santoso, M.Kom
NIP. 19770103 201101 1 004

Anggota Penguji III : Khadijah Fahmi Hayati Holle, M.Kom
NIP. 19900626 202203 2 002

()
()
()
()

Mengetahui dan Mengesahkan,
Ketua Program Studi Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Maulana Malik Ibrahim Malang


Supriyono, M. Kom
NIP. 19841010 201903 1 012

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Zainal Abidin
NIM : 220605110025
Fakultas / Program Studi : Sains dan Teknologi / Teknik Informatika
Judul Skripsi : Klasifikasi Acute Lymphoblastic Leukemia (ALL)
Menggunakan SqueezeNet Berdasarkan Citra
Sel Darah

- Menyatakan dengan sebenarnya bahwa Skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambil alihan data, tulisan, atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar pustaka.

Apabila dikemudian hari terbukti atau dapat dibuktikan skripsi ini merupakan hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan tersebut.

Malang, 22 Juni 2026
Yang membuat pernyataan,



Zainal Abidin
NIM.220605110025

MOTTO

... Dibalik kesusahan pasti ada kemudahan ...

HALAMAN PERSEMBAHAN

Alhamdulillah rabbil ‘alamin, segala puji bagi Allah SWT atas rahmat, kekuatan, dan kemudahan yang diberikan-Nya sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Shalawat dan salam senantiasa tercurahkan kepada Rasulullah Muhammad SAW yang telah menuntun umat manusia dari zaman jahiliyah menuju cahaya addinul Islam.

Dengan penuh rasa syukur, penulis mempersembahkan karya sederhana ini kepada kedua orang tua tercinta, diri sendiri yang telah berjuang hingga titik ini, adik, para dosen pembimbing, sahabat-sahabat, serta semua pihak yang telah memberikan bantuan, doa, dan semangat selama proses penyusunan skripsi ini.

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabarakatuh

Alhamdulillahirabbil'alamin, segala puji syukur penulis panjatkan ke hadirat Allah SWT. atas limpahan rahmat, nikmat, dan hidayah-Nya, sehingga penulis dapat menyelesaikan tugas skripsi yang berjudul “KLASIFIKASI *ACUTE LYMPHOBLASTIC LEUKEMIA* (ALL) MENGGUNAKAN *SQUEEZENET* BERDASARKAN CITRA SEL DARAH” dengan baik dan lancar. Shalawat serta salam tak lupa penulis sampaikan kepada Nabi Muhammad Shallallahu 'Alaihi Wasallam, sosok teladan mulia yang telah menuntun kita dari zaman kegelapan menuju zaman kebenaran, Islam, dan ilmu pengetahuan yang kita rasakan manfaatnya hingga hari ini.

Tugas skripsi ini disusun dalam rangka memenuhi salah satu syarat kelulusan untuk memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, UIN Maulana Malik Ibrahim Malang. Dalam proses penyusunan tugas ini, penulis mendapatkan bantuan, dukungan, serta doa dari banyak pihak, baik secara langsung maupun tidak langsung. Untuk itu, penulis ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada:

1. Prof. Dr. Hj. Ilfi Nur Diana, M.Si, selaku Rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang
2. Dr. H. Agus Mulyono, S.Pd., M.Kes, selaku Dekan Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim Malang.
3. Supriyono, M.Kom., selaku Ketua Program Studi S1 Teknik Informatika UIN Maulana Malik Ibrahim Malang.

4. Dr. Irwan Budi Santoso, M.Kom, selaku dosen pembimbing 1 serta mentor yang telah sangat sabar memberikan bimbingan, saran, dan dukungan yang sangat berarti bagi penulis selama proses penelitian
5. Khadijah Fahmi Hayati Holle, M.Kom, selaku Dosen Pembimbing II, yang senantiasa memberikan arahan dan masukan yang sangat berharga dalam pengerjaan skripsi ini.
6. Prof. Dr. Suhartono, S.Si, M.Kom, selaku dosen Penguji I dan Okta Qomaruddin Aziz, M.Kom, selaku dosen Penguji II yang telah memberikan kritik dan saran yang membangun dalam penyusunan skripsi ini.
7. Segenap dosen, staf administrasi, laboran, dan teman-teman mahasiswa Program Studi Teknik Informatika Fakultas Sains dan Teknologi yang telah memberikan arahan, dan dukungan yang telah diberikan kepada penulis selama proses perkuliahan dan penyusunan skripsi.

Penulis menyadari bahwa skripsi ini masih jauh dari sempurna. Oleh karena itu, segala kritik, saran, dan masukan yang membangun sangat penulis harapkan demi kesempurnaan di masa mendatang. Semoga skripsi ini bermanfaat bagi penulis khususnya, dan bagi pembaca pada umumnya.

Wassalamu'alaikum Warahmatullahi Wabarakatuh

Malang, 22 Juni 2026

Penulis

DAFTAR ISI

HALAMAN PENGAJUAN	ii
HALAMAN PERSETUJUAN	iii
HALAMAN PENGESAHAN	iv
PERNYATAAN KEASLIAN TULISAN	v
MOTTO	vi
HALAMAN PERSEMBAHAN	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
ABSTRAK	xiv
ABSTRACT	xv
البحث مستخلص	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	5
BAB II STUDI PUSTAKA	6
2.1 Penelitian Terkait	6
2.2 Acute Lymphoblastic Leukemia (ALL)	10
2.3 Klasifikasi Citra Sel Darah	13
2.4 SqueezeNet	15
BAB III DESAIN DAN IMPLEMENTASI	19
3.1 Dataset	19
3.2 Augmentasi Data	21
3.2.1 Rotasi	22
3.2.2 Flipping	24
3.3 Desain Sistem	25
3.3.1 Pra-Pemrosesan Data	26
3.3.2 Arsitektur Model SqueezeNet	29
3.4 Evaluasi Kinerja Model	39
BAB IV UJI COBA DAN PEMBAHASAN	42
4.1 Skenario Uji Coba	42
4.2 Hasil Uji Coba	44
4.2.1 Hasil Uji Coba Skenario 1	44
4.2.2 Hasil Uji Coba Skenario 2	53
4.2.3 Hasil Uji Coba Skenario 3	62
4.3 Pembahasan	73

4.3.1 Pengaruh Rasio Pembagian Data	73
4.3.2 Pengaruh Batch Size	77
4.3.3 Hasil Terbaik Keseluruhan	82
4.3.4 Perbandingan dengan Penelitian dan Arsitektur Terkait	85
BAB V KESIMPULAN DAN SARAN	88
5.1 Kesimpulan	88
5.2 Saran.....	88
DAFTAR PUSTAKA	

DAFTAR GAMBAR

Gambar 3. 1 Citra sel darah (a) Acute Lymphoblastic Leukemia (ALL) dan (b) sel darah normal (HEM).....	21
Gambar 3. 2 Hasil Rotasi Citra: (a) citra asli, (b) citra hasil rotasi.....	23
Gambar 3. 3 Hasil flipping horizontal: (a) citra asli, (b) citra hasil flipping horizontal	24
Gambar 3. 4 Desain Sistem.....	26
Gambar 3. 5 Gambar Hasil Resize.....	27
Gambar 3. 6 Normalisasi.....	28
Gambar 3. 7 Arsitektur Model SqueezeNet	30
Gambar 4. 1 Grafik Pelatihan Model Skenario 1 Batch Size 16.....	45
Gambar 4. 2 Confusion Matrix Skenario 1 Batch Size 16.....	46
Gambar 4. 3 Grafik Pelatihan Model Skenario 1 Batch Size 32.....	47
Gambar 4. 4 Confusion Matrix Skenario 1 Batch Size 32.....	48
Gambar 4. 5 Grafik Pelatihan Model Skenario 1 Batch Size 64.....	50
Gambar 4. 6 Confusion Matrix Skenario 1 Batch Size 64.....	51
Gambar 4. 7 Grafik Pelatihan Model Skenario 2 Batch Size 16.....	54
Gambar 4. 8 Confusion Matrix Skenario 2 Batch Size 16.....	55
Gambar 4. 9 Grafik Pelatihan Model Skenario 2 Batch Size 32.....	56
Gambar 4. 10 Confusion Matrix Skenario 2 Batch Size 32.....	57
Gambar 4. 11 Grafik Pelatihan Model Skenario 2 Batch Size 64.....	59
Gambar 4. 12 Confusion Matrix Skenario 2 Batch Size 64.....	60
Gambar 4. 13 Grafik Pelatihan Model Skenario 3 Batch Size 16.....	63
Gambar 4. 14 Confusion Matrix Skenario 3 Batch Size 16.....	64
Gambar 4. 15 Grafik Pelatihan Model Skenario 3 Batch Size 32.....	66
Gambar 4. 16 Confusion Matrix Skenario 3 Batch Size 32.....	67
Gambar 4. 17 Grafik Pelatihan Model Skenario 3 Batch Size 64.....	69
Gambar 4. 18 Confusion Matrix Skenario 3 Batch Size 64.....	70
Gambar 4. 19 Grafik Perbandingan Rata-rata Akurasi per Skenario.....	74
Gambar 4. 20 Grafik Perbandingan Rata-rata Akurasi dan F1-Score per Batch Size	78

DAFTAR TABEL

Tabel 2. 1 Penelitian Terdahulu dengan Machine Learning	10
Tabel 3. 1 Pembagian Dataset C-NMC 2019.....	20
Tabel 3. 2 Distribusi Kelas Dataset.....	20
Tabel 3. 3 Spesifikasi Arsitektur Model SqueezeNet	30
Tabel 3. 4 Confusion Matrix Klasifikasi Citra Sel Darah.....	39
Tabel 4. 1 Skenario Pembagian Data	42
Tabel 4. 2 Jumlah Data Latih Kelas HEM Setelah Augmentasi	43
Tabel 4. 3 Skenario Uji Coba.....	43
Tabel 4. 4 Perbandingan Hasil Evaluasi Skenario 1 (70:30)	52
Tabel 4. 5 Perbandingan Hasil Evaluasi Skenario 2 (80:20)	61
Tabel 4. 6 Perbandingan Hasil Evaluasi Skenario 3 (90:10)	72
Tabel 4. 7 Rata-rata Akurasi dan F1-Score per Skenario.....	74
Tabel 4. 8 Rata-rata Akurasi dan F1-Score per Batch Size.....	77
Tabel 4. 9 Perbandingan Hasil Seluruh Skenario dan Batch Size	83
Tabel 4. 10 Perbandingan dengan Model lain.....	85

ABSTRAK

Abidin, Zainal. 2026. **Klasifikasi *Acute Lymphoblastic Leukemia* (ALL) Menggunakan SqueezeNet Berdasarkan Citra Sel Darah.** Skripsi. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing: (I) Dr. Irwan Budi Santoso, M.Kom (II) Khadijah Fahmi Hayati Holle, M.Kom.

Kata kunci: Leukemia, SqueezeNet, Convolutional Neural Network, Klasifikasi Citra, Deep Learning.

Leukemia adalah salah satu jenis kanker darah yang memerlukan diagnosis dini yang cepat dan akurat. Diagnosis manual melalui pemeriksaan mikroskopis sel darah membutuhkan keahlian khusus dan waktu yang lama, sehingga diperlukan pendekatan otomatis berbasis kecerdasan buatan untuk membantu proses klasifikasi. Penelitian ini menggunakan model SqueezeNet yang dilengkapi dengan 10 unit Fire untuk mengklasifikasikan gambar sel darah terkait leukemia limfoblastik akut (ALL) dan sel darah normal (HEM) pada dataset C-NMC 2019. Evaluasi dilakukan pada tiga skenario pembagian data (70:30, 80:20, 90:10) dan tiga ukuran batch yang berbeda (16, 32, 64). Hasil terbaik diperoleh pada skenario 80:20 dengan ukuran batch 16, di mana akurasi mencapai 84,36%, akurasi spesifik 84,5%, tingkat pemulihan 79%, dan indeks F1 80,5%. Studi ini juga menyimpulkan bahwa ukuran batch optimal bersifat dinamis — batch kecil lebih efektif pada data pelatihan yang terbatas, sedangkan batch besar memberikan estimasi gradien yang lebih stabil pada data pelatihan yang besar. Hasil ini menunjukkan bahwa SqueezeNet memiliki potensi sebagai model ringan dengan kinerja klasifikasi yang baik, sehingga dapat dikembangkan sebagai teknologi pendukung untuk diagnosis leukemia di lingkungan medis yang memiliki keterbatasan sumber daya komputasi.

ABSTRACT

Abidin, Zainal. 2026. **Classification of Acute Lymphoblastic Leukemia (ALL) Using SqueezeNet Based on Blood Cell Images.** Thesis. Informatics Engineering Study Program, Faculty of Science and Technology, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors: (I) Dr. Irwan Budi Santoso, M.Kom (II) Khadijah Fahmi Hayati Holle, M.Kom.

Key words: *Leukemia, SqueezeNet, Convolutional Neural Network, Image Classification, Deep Learning.*

Leukemia is a type of blood cancer that requires rapid and accurate early diagnosis. Manual diagnosis through microscopic examination of blood cells requires specialized expertise and a significant amount of time, making an automated, artificial intelligence-based approach necessary to assist in the classification process. This study uses a SqueezeNet model equipped with 10 Fire units to classify images of blood cells associated with acute lymphoblastic leukemia (ALL) and normal blood cells (HEM) in the C-NMC 2019 dataset. Evaluation was conducted across three data split scenarios (70:30, 80:20, 90:10) and three different batch sizes (16, 32, 64). The best results were obtained in the 80:20 scenario with a batch size of 16, where accuracy reached 84.36%, specific accuracy 84.5%, recall 79%, and F1 score 80.5%. This study also concluded that the optimal batch size is dynamic smaller batches are more effective with limited training data, while larger batches provide more stable gradient estimates with large training datasets. These results indicate that SqueezeNet has the potential to serve as a lightweight model with good classification performance, making it suitable for development as a supporting technology for leukemia diagnosis in medical settings with limited computational resources.

مستخلص البحث

عابدين، زينال. 2026. تصنيف سرطان الدم الليمفاوي الحاد (ALL) باستخدام SqueezeNet استنادًا إلى صور خلايا الدم. أطروحة البكالوريوس. قسم هندسة المعلوماتية، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية في مالانغ. المشرفان (I): د. إيروان بودي سانتوسو، ماجستير في علوم الحاسوب (II) خديجة فهمي حياة هول، ماجستير في علوم الحاسوب.

الكلمات المفتاحية: اللوكيميا، SqueezeNet، الشبكة العصبية التلافيفية، تصنيف الصور، التعلم العميق.

اللوكيميا هي أحد أنواع سرطان الدم التي تتطلب تشخيصًا مبكرًا وسريعًا ودقيقًا. يتطلب التشخيص اليدوي من خلال الفحص المجهرى لخلايا الدم خبرة خاصة ووقتًا طويلاً، لذا فإن هناك حاجة إلى نهج آلي قائم على الذكاء الاصطناعي للمساعدة في عملية التصنيف. تستخدم هذه الدراسة نموذج SqueezeNet المزود بـ 10 وحدات Fire لتصنيف صور خلايا الدم الخاصة بسرطان الدم الليمفاوي الحاد (ALL) وخلايا الدم الطبيعية (HEM) في مجموعة البيانات C-NMC 2019. أُجري التقييم على ثلاثة سيناريوهات لتقسيم البيانات (70:30، 80:20، 90:10) وثلاثة أحجام مختلفة للدفعات (16، 32، 64). وقد تم الحصول على أفضل النتائج في سيناريو 80:20 مع حجم الدفعة 16، حيث بلغت الدقة 84,36%، والدقة النوعية 84,5%، ومعدل الاسترجاع 79%، ومؤشر F1 80,5%. كما توصلت هذه الدراسة إلى أن الحجم الأمثل للدفعة ديناميكي — فالدفعات الصغيرة تكون أكثر فعالية في حالة البيانات التدريبية المحدودة، في حين توفر الدفعات الكبيرة تقديرات أكثر استقرارًا للتدرج في حالة البيانات التدريبية الكبيرة. وتشير هذه النتائج إلى أن SqueezeNet يمتلك إمكانيات كونه نموذجًا خفيًا يتمتع بأداء تصنيفي جيد، مما يجعله قابلاً للتطوير كتقنية داعمة لتشخيص سرطان الدم في البيئات الطبية التي تعاني من محدودية موارد الحوسبة.

BAB I

PENDAHULUAN

1.1 Latar Belakang

Leukemia merupakan salah satu jenis kanker darah yang ditandai oleh bertambahnya sel darah putih abnormal secara cepat dan tidak terkendali di dalam sumsum tulang. Keadaan ini mengganggu tahapan pembentukan sel darah sehat yang berfungsi membawa oksigen, melawan infeksi, dan menghentikan perdarahan, sehingga pasien sangat rentan mengalami anemia berat, infeksi berulang, serta perdarahan spontan. Jenis leukemia seperti Acute Lymphoblastic Leukemia (ALL) bersifat sangat agresif dan dapat memburuk dalam waktu singkat apabila tidak ditangani secara cepat. Penelitian oleh Zhang (2023) menegaskan bahwa keterlambatan dalam mendeteksi sel abnormal ialah salah satu faktor penyebab tingginya angka kematian pada pasien leukemia. Oleh karena itu, diperlukan teknologi pendukung diagnosis yang mampu menjalankan klasifikasi sel darah penderita leukemia dengan cepat, akurat, dan tetap ringan sehingga dapat meningkatkan peluang keberhasilan penanganan medis.

Dalam perspektif Islam, pengembangan ilmu pengetahuan dan teknologi merupakan bagian dari usaha manusia dalam mengerti tanda-tanda kebesaran Allah SWT melalui pengamatan dan penelitian terhadap ciptaan-Nya. Pemanfaatan teknologi dalam bidang kesehatan, termasuk dalam analisis citra medis, dapat menjadi salah satu bentuk ikhtiar untuk membantu tahapan diagnosis penyakit

dengan lebih cepat dan akurat. Hal ini sejalan dengan firman Allah SWT dalam (QS. Yunus [10]: 101):

قُلْ انظُرُوا مَاذَا فِي السَّمٰوٰتِ وَالْاَرْضِ يَوْمَ تُغْنِ الْاٰيٰتُ وَالنُّذُرُ عَنْ قَوْمٍ لَا يُؤْمِنُوْنَ ﴿١٠١﴾

“Katakanlah: Perhatikanlah apa yang ada di langit dan di bumi. Tidaklah bermanfaat tanda-tanda (kebesaran Allah) dan rasul-rasul yang memberi peringatan bagi orang-orang yang tidak beriman”(Q.S Yunus [10] : 101)

Berdasarkan Shihab (2002) dalam Tafsir Al-Misbah, kata "perhatikanlah" pada ayat tersebut bermakna memperhatikan dengan mata kepala sekaligus dengan hati, yakni menggunakan akal yang dianugerahkan Allah SWT untuk mengamati makhluk dan sistem kerja yang ada di langit dan di bumi. Beliau menegaskan bahwa penggunaan akal secara optimal dalam mengamati alam raya sudah cukup untuk mengantarkan manusia pada kesadaran akan kekuasaan dan keesaan Allah SWT. Dalam konteks perkembangan ilmu pengetahuan modern, perintah untuk memperhatikan dan mengamati tersebut dapat diwujudkan melalui kegiatan penelitian dan kemajuan teknologi yang memberi manfaat bagi kehidupan manusia.

Dalam penelitian ini, proses analisis citra sel darah menggunakan teknologi deep learning merupakan salah satu bentuk implementasi ilmu pengetahuan dalam bidang kesehatan. Penggunaan model SqueezeNet untuk mengklasifikasikan citra sel darah Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM) dilakukan melalui proses pengamatan pola visual pada citra medis secara sistematis dan objektif. Dengan begitu, studi ini diinginkan mampu mendukung pengembangan teknologi diagnosis berbasis citra yang lebih cepat, akurat, dan efisien sebagai bentuk pemanfaatan ilmu pengetahuan untuk kemaslahatan manusia.

Berbagai penelitian sebelumnya telah menerapkan metode deep learning dalam klasifikasi citra sel darah untuk mendeteksi leukemia. Ahmed et al (2020) memanfaatkan VGGNet dengan optimasi berbasis swarm untuk meningkatkan akurasi klasifikasi sel darah putih. Mulya *et al* (2026) menunjukkan bahwa model *Convolutional Neural Network* berbasis InceptionV3 mampu mengklasifikasikan citra sel darah putih untuk mendeteksi Acute Lymphoblastic Leukemia (ALL) dengan tingkat akurasi yang tinggi. Zhang (2023) mengoptimalkan arsitektur CNN untuk mendeteksi ALL pada citra grayscale. Akbar et al (2024) mengemukakan bahwa penerapan Convolutional Neural Network pada citra mikroskopis sel darah dapat digunakan secara efektif sebagai alat bantu deteksi leukemia limfoblastik akut. Sampathila *et al* (2022) mengembangkan arsitektur CNN kustom untuk mendeteksi Acute Lymphoblastic Leukemia dari citra apusan darah, yang menunjukkan bahwasanya pendekatan deep learning mampu menumbuhkan akurasi diagnosis otomatis. Hal ini menunjukkan bahwasanya sebagian besar penelitian masih fokus pada arsitektur CNN berukuran besar, sementara pemanfaatan arsitektur ringan seperti SqueezeNet pada domain citra leukemia masih relatif terbatas. Padahal, Iandola et al (2017) menyebutkan bahwa SqueezeNet memiliki keunggulan berupa ukuran model yang sangat kecil yang kurang dari 5 MB, kecepatan inferensi yang tinggi, serta keperluan komputasi yang rendah, sehingga sangat cocok untuk diimplementasikan pada perangkat medis dengan sumber daya terbatas, seperti sistem point-of-care, perangkat portabel, maupun integrasi ke aplikasi klinis berbasis edge computing. Keunggulan tersebut membuat SqueezeNet potensial untuk digunakan dalam klasifikasi leukemia yang

cepat dan efisien, meskipun belum banyak penelitian yang mengevaluasi performanya secara mendalam pada citra mikroskopis sel darah.

Kondisi tersebut membuka sebuah *gap penelitian* yang jelas karena belum banyak penelitian yang secara khusus memfokuskan diri pada penerapan dan peningkatan performa SqueezeNet dalam klasifikasi leukemia berbasis citra sel darah, meskipun model ini memiliki karakteristik yang ideal untuk lingkungan medis. Dengan demikian, diperlukan penelitian yang secara khusus mengimplementasikan dan mengevaluasi SqueezeNet pada dataset leukemia, sehingga model yang ringan ini dapat dimanfaatkan secara optimal untuk kebutuhan diagnosa awal. Berdasarkan kebutuhan tersebut, penelitian ini mengusulkan penggunaan SqueezeNet untuk melakukan klasifikasi citra sel darah pada dataset leukemia guna mengevaluasi performa model ringan ini dalam tugas diagnosis berbasis citra. Pendekatan ini diinginkan tidak hanya memberi kontribusi akademik terhadap pengembangan model deep learning yang efisien, namun juga mendukung implementasi teknologi diagnostik yang cepat, mudah digunakan, dan hemat komputasi bagi lingkungan medis.

1.2 Rumusan Masalah

Rumusan masalah dalam ini adalah bagaimana performa model SqueezeNet dalam mengklasifikasikan citra mikroskopis sel darah penderita Acute Lymphoblastic Leukemia (ALL) secara akurat sebagai pendukung diagnosis berbasis citra.

1.3 Batasan Masalah

1. Studi ini memanfaatkan dataset citra mikroskopis sel darah yang didapatkan dari Kaggle, yakni dataset Leukemia Classification (C-NMC 2019) yang disediakan oleh (Gupta dan Gupta, 2019)
2. Ruang lingkup penelitian difokuskan pada klasifikasi citra sel darah normal (HEM) dan sel darah penderita Acute Lymphoblastic Leukemia (ALL).

1.4 Tujuan Penelitian

Tujuan dari studi ini ialah untuk mengevaluasi performa model SqueezeNet dalam mengklasifikasikan citra mikroskopis sel darah normal (HEM) dan sel darah penderita Acute Lymphoblastic Leukemia (ALL) menggunakan dataset C-NMC 2019 berdasarkan metrik accuracy, recall, precision, dan F1-score, sehingga mampu dimanfaatkan sebagai pendukung diagnosis Acute Lymphoblastic Leukemia (ALL) berbasis citra sel darah.

1.5 Manfaat Penelitian

Studi ini diinginkan mampu memberi manfaat bagi tenaga medis dan rumah sakit sebagai pendukung diagnosis Acute Lymphoblastic Leukemia (ALL) melewati klasifikasi citra mikroskopis sel darah, sehingga membantu pengambilan keputusan medis secara lebih cepat, tepat, dan efisien.

BAB II

STUDI PUSTAKA

2.1 Penelitian Terkait

Penelitian terkait klasifikasi citra sel darah leukemia berbasis *deep learning* menunjukkan perkembangan yang signifikan dalam beberapa tahun terakhir, seiring dengan naiknya keperluan akan sistem diagnosis yang objektif, konsisten, dan mampu mendukung pengambilan keputusan medis. Leukemia, khususnya *Acute Lymphoblastic Leukemia* (ALL), memiliki karakteristik morfologi sel yang kompleks dan sering kali sulit dibedakan secara visual, terutama pada tahap awal penyakit. Oleh karena itu, pemanfaatan citra mikroskopis *blood smear* yang dianalisis menggunakan pendekatan komputasi menjadi salah satu solusi yang banyak dikaji.

Teknik berbasis Convolutional Neural Network (CNN) mendominasi penelitian klasifikasi citra sel darah leukemia sebab kemampuan ekstraksi fitur visual yang dimilikinya dapat bekerja secara otomatis langsung dari citra, sehingga mengurangi ketergantungan terhadap proses rekayasa fitur yang dilakukan secara manual. CNN mampu mempelajari karakteristik penting seperti bentuk inti sel, ukuran sel, tekstur sitoplasma, serta distribusi intensitas warna tanpa membutuhkan proses ekstraksi fitur manual. Perihal ini menjadikan CNN sangat sesuai untuk mengatasi kompleksitas citra medis yang memiliki variasi morfologi tinggi.

Dalam konteks ketersediaan data, Gupta & Gupta (2019) memperkenalkan dataset C-NMC 2019 yang berisi citra mikroskopis sel darah dengan dua kelas

utama, yaitu *Acute Lymphoblastic Leukemia* (ALL) dan sel darah normal (HEM). Dataset ini telah dianotasi oleh ahli hematologi sehingga kualitas label data relatif terjamin dan banyak digunakan sebagai dataset acuan dalam penelitian klasifikasi leukemia berbasis citra. Relevansi dan konsistensi dataset ini kembali ditegaskan oleh Gupta *et al* (2022), yang menyatakan bahwa C-NMC 2019 masih layak digunakan sebagai *benchmark* dalam evaluasi performa model klasifikasi ALL.

Sejumlah penelitian memanfaatkan arsitektur CNN berukuran besar untuk mencapai tingkat akurasi yang tinggi. Ullah *et al* (2021) mengembangkan model CNN berbasis VGG untuk membedakan sel darah leukemik dan sel darah normal. Meskipun arsitektur VGG memperlihatkan keterampilan yang baik dalam mengekstraksi fitur visual sel darah, penggunaan arsitektur ini memiliki keterbatasan berupa jumlah parameter yang besar yang berdampak pada tingginya keperluan sumber daya komputasi. Meskipun menghasilkan akurasi yang baik, penggunaan arsitektur tersebut masih memiliki keterbatasan dari sisi efisiensi model.

Pendekatan *deep CNN* dengan struktur jaringan yang lebih dalam juga banyak diterapkan dalam penelitian klasifikasi leukemia. Huang & Huang (2024) menggunakan *deep CNN* pada citra *blood smear* leukemia dan melaporkan performa klasifikasi yang tinggi. Namun, penelitian tersebut menunjukkan bahwa peningkatan performa dicapai dengan konsekuensi meningkatnya kompleksitas komputasi dan waktu pelatihan model. Hal ini menunjukkan bahwa penggunaan jaringan CNN berukuran besar masih menghadapi kendala ketika diterapkan pada sistem dengan keterbatasan sumber daya.

Dalam upaya meningkatkan performa klasifikasi, beberapa penelitian mengombinasikan atau membandingkan berbagai arsitektur CNN modern. Mondal *et al* (2021) menerapkan pendekatan *ensemble CNN* dengan menggabungkan beberapa arsitektur, seperti GoogLeNet, ResNet, VGG, MobileNet, Inception, DenseNet, dan AlexNet pada dataset C-NMC 2019. Pendekatan ini mampu meningkatkan stabilitas dan akurasi klasifikasi, namun juga menyebabkan peningkatan kompleksitas sistem secara signifikan. Studi komparatif serupa dilakukan oleh Fajrina *et al* (2024) yang membandingkan berbagai arsitektur CNN modern dan menunjukkan bahwa arsitektur dengan kompleksitas tinggi cenderung menghasilkan performa yang lebih baik, namun penerapannya memerlukan kebutuhan sumber daya komputasi yang tidak sedikit.

Penelitian lain oleh Vinnarasi *et al* (2024) juga menerapkan *deep CNN* pada citra *blood smear* ALL dan memperlihatkan bahwasanya CNN mampu mempelajari fitur visual sel darah dengan cara efektif tanpa ekstraksi fitur manual. Meskipun demikian, kedua penelitian tersebut masih mengandalkan model dengan jumlah parameter yang relatif besar, sehingga aspek efisiensi model belum menjadi fokus utama.

Penelitian di tingkat nasional juga menunjukkan tren penggunaan CNN berukuran besar. Istiqomah & Ardianto (2024) mengembangkan CNN konvensional dengan arsitektur kustom untuk klasifikasi leukemia dan melaporkan performa yang baik, namun masih menghadapi tantangan terkait efisiensi komputasi. Selain itu, Alifian *et al* (2024) mengusulkan penggunaan arsitektur EfficientNetV2M yang dioptimasi menggunakan Bayesian Optimization untuk

klasifikasi leukemia. Pendekatan ini berhasil meningkatkan akurasi klasifikasi, tetapi memiliki kompleksitas jaringan yang tinggi dan kebutuhan komputasi yang besar.

Selain itu, Zhang (2023) mengembangkan CNN kustom berbasis citra grayscale untuk klasifikasi ALL dan menunjukkan bahwa penyesuaian *hyperparameter* dapat meningkatkan performa klasifikasi. Meskipun pendekatan ini efektif, arsitektur yang digunakan masih tergolong CNN konvensional dengan ukuran model yang tidak kecil, sehingga efisiensi model tetap menjadi tantangan.

Berdasarkan kajian terhadap berbagai penelitian yang telah dilakukan, CNN terbukti menjadi pendekatan yang efektif dalam klasifikasi citra sel darah leukemia, terutama dalam membedakan sel ALL dan HEM. Akan tetapi, fokus utama sebagian besar penelitian tersebut masih tertuju pada peningkatan akurasi melalui penggunaan arsitektur CNN yang kompleks dengan jumlah parameter yang besar, sehingga aspek efisiensi model belum mendapat perhatian yang memadai. Aspek efisiensi model, seperti ukuran jaringan, jumlah parameter, kebutuhan memori, dan kecepatan inferensi, belum menjadi fokus utama dalam banyak penelitian.

Oleh karena itu, masih terdapat celah penelitian dalam pengembangan sistem klasifikasi citra sel darah leukemia yang tak hanya akurat, namun juga efisien dengan cara komputasi. Penelitian ini memfokuskan pada penggunaan arsitektur CNN ringan, yaitu SqueezeNet, untuk melakukan klasifikasi citra sel darah menjadi dua kelas, yakni *Acute Lymphoblastic Leukemia* (ALL) dan sel darah normal (HEM). Pemilihan SqueezeNet didasarkan pada karakteristiknya yang memiliki jumlah parameter jauh lebih kecil dibandingkan arsitektur CNN pada

umumnya, namun kemampuan ekstraksi fitur visualnya tetap dapat berjalan secara efektif. Maka dari itu, studi ini diinginkan mampu memberi kontribusi nyata dalam pengembangan sistem klasifikasi citra sel darah leukemia yang dapat diterapkan pada lingkungan medis yang memiliki keterbatasan sumber daya komputasi.

Tabel 2. 1 Penelitian Terdahulu dengan Machine Learning

No	Peneliti	Metode	Dataset	Hasil
1	Mondal et al. (2021)	Ensemble CNN (GoogLeNet, ResNet, VGG, MobileNet, Inception, DenseNet, AlexNet)	C-NMC 2019 (ALL & HEM)	Akurasi = 81–93%
2	Gupta & Gupta (2022)	CNN berbasis VGG, AlexNet dan - ResNet	C-NMC 2019 (ALL & HEM)	Akurasi = 84-95%
3	Huang et al. (2021)	AlexNet	C-NMC 2019 (ALL & HEM)	Akurasi = 84%
4	Fajrina et al. (2023)	EfficientNet	C-NMC 2019 (ALL & HEM)	Akurasi = 96%
5	Ullah et al. (2021)	CNN berbasis VGG	C-NMC 2019 (ALL & HEM)	Akurasi = 90%
6	Khan et al. (2023)	Deep CNN	Citra sel darah leukemia	Akurasi = 96%
7	Istiqomah et al. (2024)	CNN konvensional (custom architecture)	Citra leukemia	Akurasi = 94%
8	Vinnarasi et al. (2024)	Deep CNN	Blood smear ALL	Akurasi = 91%
9	Wibowo et al. (2024)	EfficientNetV2M + Bayesian Optimization	Citra leukemia	Akurasi = 91%
10	Zhang et al. (2024)	CNN kustom (grayscale images)	Citra ALL	Akurasi = 80%

2.2 Acute Lymphoblastic Leukemia (ALL)

Acute Lymphoblastic Leukemia (ALL) ialah salah satu macam leukemia akut yang ditandai oleh proliferasi sel limfoid imatur atau sel blast dengan cara tidak terkontrol di dalam sumsum tulang dan darah perifer. Akumulasi sel blast tersebut menyebabkan terganggunya proses pembentukan sel darah normal, termasuk leukosit matang, eritrosit, dan trombosit, sehingga penderita ALL sering mengalami anemia, infeksi berulang, serta gangguan perdarahan. Chang *et al* (2021) menyebutkan bahwa ALL ialah jenis leukemia yang paling sering dijumpai pada

anak-anak, meskipun insidensinya juga ditemukan pada kelompok remaja dan dewasa dengan karakteristik klinis yang berbeda.

Secara biologis, ALL berasal dari prekursor limfosit B atau limfosit T yang mengalami gangguan proses diferensiasi dan maturasi. Inaba & Pui (2021) menjelaskan bahwa kegagalan pematangan sel limfoid tersebut menyebabkan sel blast terus bereplikasi tanpa kendali dan mendominasi sumsum tulang. Kondisi ini tidak hanya menghambat produksi sel darah normal, tetapi juga menyebabkan penyebaran sel blast ke darah perifer dan jaringan lain.

Dari sisi morfologi, sel blast ALL memiliki ciri khas yang dapat diamati melalui pemeriksaan mikroskopis. Sekar & Raj (2023) menjelaskan bahwa sel blast ALL umumnya memiliki inti sel berukuran besar dengan rasio inti terhadap sitoplasma yang tinggi, kromatin yang tampak padat atau halus tidak merata, serta sitoplasma yang relatif sempit. Selain itu, perbedaan intensitas warna dan tekstur antara inti sel dan sitoplasma menjadi karakteristik visual yang membedakan sel ALL dari sel darah normal. Karakteristik morfologi ini menjadi dasar penting dalam analisis citra sel darah.

Dalam praktik klinis, diagnosis ALL dilakukan melalui kombinasi pemeriksaan laboratorium, termasuk pemeriksaan darah tepi, aspirasi sumsum tulang, serta analisis morfologi sel secara mikroskopis. Namun, pendekatan konvensional ini sangat bergantung pada keahlian dan pengalaman tenaga medis, sehingga berpotensi menimbulkan variasi interpretasi antar-pemeriksa. Chang *et al* (2021) menegaskan bahwa perbedaan pengalaman analisis laboratorium dapat

memengaruhi konsistensi hasil identifikasi sel blast, terutama pada kasus dengan perbedaan morfologi yang tidak terlalu mencolok.

Perkembangan teknologi pengolahan citra dan kecerdasan buatan memberikan peluang untuk memberi dukungan analisis sel darah penderita ALL secara lebih objektif. Elsayed *et al* (2023) menjelaskan bahwa citra mikroskopis sel darah mengandung informasi visual yang kaya, seperti pola tekstur, bentuk sel, dan distribusi warna, yang dapat dianalisis menggunakan pendekatan komputasional. Dengan memanfaatkan teknik deep learning, karakteristik visual sel blast ALL dapat dipelajari secara otomatis dari data citra tanpa bergantung sepenuhnya pada penilaian subjektif manusia.

Pemahaman yang baik mengenai karakteristik biologis dan morfologi sel ALL menjadi aspek krusial dalam pengembangan sistem klasifikasi citra sel darah. Inaba & Pui (2021) menyatakan bahwa perbedaan struktur inti sel, variasi tekstur kromatin, serta perubahan morfologi sitoplasma merupakan fitur visual utama yang dapat dimanfaatkan dalam sistem klasifikasi berbasis citra. Oleh karena itu, kajian mengenai ALL tidak hanya berperan sebagai landasan medis, namun juga sebagai dasar teoritis dalam penelitian pengolahan citra sel darah.

Selain itu, variasi morfologi sel blast ALL antar pasien dan antar tahap penyakit menjadi tantangan tersendiri dalam proses analisis citra. Sekar & Raj (2023) menjelaskan bahwa heterogenitas sel blast dapat memengaruhi tingkat kesulitan klasifikasi, sehingga dibutuhkan pendekatan yang mampu menangkap pola visual kompleks dengan cara konsisten. Hal ini memperkuat urgensi

penggunaan metode berbasis deep learning yang memiliki kemampuan representasi fitur yang baik.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa Acute Lymphoblastic Leukemia (ALL) merupakan penyakit dengan karakteristik morfologi sel yang khas dan dapat dianalisis melalui citra mikroskopis. Pemahaman mendalam mengenai ciri biologis dan visual sel blast ALL sangat penting sebagai landasan teoritis dalam penelitian ini. Kajian ini menjadi dasar bagi pengembangan sistem klasifikasi citra sel darah yang berfokus pada pengelompokan sel Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM) sebagai bagian dari pendukung diagnosis leukemia berbasis citra.

2.3 Klasifikasi Citra Sel Darah

Klasifikasi citra sel darah merupakan salah satu pendekatan penting dalam pemanfaatan teknologi pengolahan citra dan kecerdasan buatan di bidang medis, khususnya dalam analisis penyakit berbasis citra mikroskopis. Proses klasifikasi tujuannya untuk mengelompokkan citra sel darah ke dalam kelas tertentu menurut karakteristik visual yang dimilikinya, seperti bentuk sel, ukuran inti, rasio inti terhadap sitoplasma, tekstur sitoplasma, serta distribusi warna. Pendekatan ini memungkinkan analisis sel darah dilakukan secara lebih objektif dan konsisten dibanding pemeriksaan manual yang sangat bergantung pada pengalaman tenaga medis.

Dalam konteks analisis sel darah putih, Tavakoli *et al* (2021) menjelaskan bahwa citra mikroskopis sel darah mengandung informasi visual yang dapat dimanfaatkan untuk membedakan jenis sel melalui analisis morfologi inti sel,

tekstur, serta fitur spasial lainnya. Penelitian tersebut menunjukkan bahwa pendekatan berbasis pengolahan citra dan pembelajaran mesin mampu meningkatkan akurasi klasifikasi sel darah secara signifikan.

Kemajuan di bidang *deep learning*, khususnya pada arsitektur Convolutional Neural Network (CNN), telah memberikan dampak signifikan terhadap peningkatan performa klasifikasi citra sel darah. (Huang & Huang, 2024) mengungkapkan bahwa CNN mempunyai keterampilan untuk mengekstraksi fitur visual dengan otomatis dari citra mikroskopis tanpa membutuhkan proses ekstraksi fitur dengan cara manual, sehingga model dapat mengenali pola-pola kompleks yang sulit diidentifikasi menggunakan pendekatan konvensional.

Pendekatan klasifikasi citra sel darah juga berperan penting dalam sistem pendukung diagnosis berbasis citra. Özcan & Uyars (2024) menegaskan bahwa sistem klasifikasi dan segmentasi sel darah putih berbasis deep learning dapat membantu mempercepat proses analisis laboratorium dengan tetap mempertahankan tingkat akurasi yang tinggi, sehingga berpotensi mengurangi beban kerja tenaga medis.

Dalam konteks data citra leukemia dan blood smear, Shahzad et al (2024) menunjukkan bahwa model deep learning mampu memanfaatkan informasi visual citra sel darah untuk membedakan sel normal dan sel abnormal secara efektif. Hasil penelitian tersebut menegaskan bahwa pendekatan klasifikasi berbasis citra memiliki potensi besar dalam mendukung analisis penyakit hematologi.

Namun demikian, tantangan utama dalam klasifikasi citra sel darah masih berkaitan dengan efisiensi model dan kebutuhan sumber daya komputasi. June et al

(2025) menyatakan bahwa walaupun CNN mampu menghasilkan performa klasifikasi yang cukup baik,

Berdasarkan beberapa penelitian tersebut, dapat disimpulkan bahwa klasifikasi citra sel darah berbasis deep learning merupakan pendekatan yang efektif dan menjanjikan dalam analisis citra mikroskopis. Pemanfaatan karakteristik visual sel darah memungkinkan pengelompokan sel secara sistematis dan objektif, sehingga klasifikasi citra sel darah menjadi landasan penting dalam pengembangan sistem pendukung diagnosis berbasis citra.

2.4 SqueezeNet

SqueezeNet mengadopsi pendekatan *fire modules* untuk meminimalkan jumlah parameter. Keunggulan ini menjadi relevan dalam konteks medical imaging, di mana deployment pada perangkat dengan resource terbatas (*point-of-care*, *edge devices*) menjadi pertimbangan praktis.

Menurut Hao et al (2024), SqueezeNet dikembangkan dengan prinsip efisiensi parameter melalui penggunaan struktur *fire module*. Modul ini terdiri atas dua bagian utama, yaitu *squeeze layer* dan *expand layer*. *Squeeze layer* berfungsi untuk mengurangi jumlah kanal fitur menggunakan konvolusi berukuran kecil, sedangkan *expand layer* bertugas memperkaya representasi fitur dengan memanfaatkan kombinasi konvolusi berukuran kecil. Pendekatan ini memungkinkan jaringan untuk mempertahankan kemampuan representasi fitur yang baik dengan jumlah parameter yang jauh lebih sedikit dibandingkan arsitektur CNN konvensional.

Minu et al (2022) menjelaskan bahwa penggunaan SqueezeNet dapat memberikan keseimbangan yang baik antara kompleksitas model dan performa klasifikasi. Dalam penelitiannya, SqueezeNet menunjukkan kemampuan untuk mempelajari pola visual penting dari citra tanpa memerlukan struktur jaringan yang terlalu dalam. Hal ini menjadi keunggulan tersendiri ketika SqueezeNet diterapkan pada dataset citra dengan karakteristik visual yang relatif homogen, seperti citra mikroskopis sel darah.

Dalam konteks penelitian berbasis citra medis, (Setiawan et al., 2021) menyatakan bahwa arsitektur CNN ringan memiliki potensi yang besar karena mampu meminimalkan risiko *overfitting*. Dengan jumlah parameter yang lebih rendah, SqueezeNet secara inherent memiliki efek regularisasi yang mengurangi *overfitting* pada dataset ALL yang terbatas. Ini berbeda dari model besar seperti VGG atau ResNet, yang lebih rentan *memorizing training samples* ketika dataset kecil.

Penelitian yang dilakukan oleh Danendra et al (2023) memperlihatkan bahwasanya CNN ringan mampu digunakan secara efektif untuk beberapa tugas klasifikasi citra dengan tingkat akurasi yang kompetitif. Studi tersebut menekankan bahwa efisiensi arsitektur tidak selalu berbanding lurus dengan penurunan performa klasifikasi, selama struktur jaringan dirancang dengan mempertimbangkan karakteristik data yang digunakan. Hal ini mendukung penggunaan SqueezeNet sebagai arsitektur yang layak untuk penelitian klasifikasi citra sel darah.

Harahap et al (2024) juga mengungkapkan bahwa arsitektur CNN ringan memiliki keunggulan dalam hal kesederhanaan struktur jaringan dan kemudahan

dalam proses pelatihan. Model dengan struktur yang lebih sederhana memungkinkan peneliti untuk lebih fokus pada analisis performa model tanpa harus menghadapi permasalahan kompleksitas arsitektur yang berlebihan. Dalam penelitian tersebut, CNN ringan mampu menghasilkan hasil klasifikasi yang konsisten pada berbagai jenis citra.

Desiana et al (2024) menambahkan bahwa CNN dengan struktur ringan sangat sesuai digunakan sebagai dasar pengembangan sistem pendukung diagnosis berbasis citra. Model yang efisien memungkinkan proses klasifikasi dilakukan secara lebih terkontrol, sehingga hasil yang diperoleh dapat dianalisis dengan lebih baik. Hal ini menjadi penting dalam konteks penelitian yang bertujuan untuk mengelompokkan citra berdasarkan karakteristik visual utama, seperti pada citra sel darah leukemia.

Selain itu, Baygin (2022) menjelaskan bahwa penggunaan CNN ringan dapat membantu menjaga keseimbangan antara kemampuan generalisasi model dan ketepatan klasifikasi. Arsitektur yang tidak terlalu kompleks cenderung lebih fokus pada fitur-fitur visual yang benar-benar relevan, sehingga dapat mengurangi pengaruh noise pada citra. Temuan ini semakin memperkuat relevansi penggunaan SqueezeNet dalam penelitian klasifikasi citra medis.

Menurut beberapa penelitian tersebut, mampu ditarik kesimpulan bahwasanya SqueezeNet merupakan arsitektur CNN yang memiliki karakteristik sesuai untuk tugas klasifikasi citra sel darah. Struktur jaringan yang sederhana, jumlah parameter yang relatif sedikit, serta keterampilan ekstraksi fitur yang memadai menjadikan SqueezeNet sebagai pilihan yang tepat untuk penelitian ini.

Oleh karena itu, SqueezeNet digunakan sebagai arsitektur utama dalam mengklasifikasikan citra sel darah menjadi dua kelas, yakni Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM), sebagai bagian dari pengembangan sistem pendukung diagnosis leukemia berbasis citra mikroskopis.

BAB III

DESAIN DAN IMPLEMENTASI

3.1 Dataset

Dataset yang digunakan dalam penelitian ini adalah dataset publik C-NMC 2019 yang dikembangkan oleh Gupta dan Gupta (2019) dan dapat diakses melalui platform Kaggle. Dataset ini dirancang untuk mendukung penelitian klasifikasi sel darah putih leukemia jenis Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM) berbasis citra mikroskopis. Label kelas pada setiap citra telah dianotasi oleh ahli onkologi sehingga kualitas *ground truth* data dapat dipertanggungjawabkan secara klinis.

Citra pada dataset C-NMC 2019 memiliki format BMP (*Bitmap*) dengan ukuran 450×450 piksel dan ruang warna RGB (3 *channel*). Setiap citra merepresentasikan satu sel darah putih yang telah disegmentasi dari preparat apusan darah mikroskopis. Citra dalam dataset ini mengandung variasi alami berupa perbedaan pewarnaan dan pencahayaan, namun telah melalui proses normalisasi warna menggunakan metode GCTI-SN yang dikembangkan oleh tim peneliti dataset.

Dataset C-NMC 2019 secara keseluruhan terdiri dari 15.135 citra yang berasal dari 118 subjek, terdiri dari 69 subjek penderita ALL dan 49 subjek normal. Dataset dibagi ke dalam tiga bagian utama sebagaimana ditunjukkan pada Tabel 3.1.

Tabel 3. 1 Pembagian Dataset C-NMC 2019

Bagian Data	Jumlah Subjek	Jumlah Citra
Training Data	73 subjek	10.661 citra
Validation Data	28 subjek	1.867 citra
Testing Data	17 subjek	2.586 citra
Total	118 subjek	15.135 citra

Pembagian data dilakukan berdasarkan subjek sehingga tidak ada subjek yang sama yang muncul di lebih dari satu bagian data. Dalam penelitian ini, data yang digunakan berasal dari penggabungan *training data* dan *validation data*, sehingga total data yang digunakan berjumlah 12.528 citra. *Testing data* tidak digunakan dalam penelitian ini karena tidak disertai label kelas, sehingga tidak dapat digunakan dalam skema *supervised learning*. Dari total 12.528 citra yang digunakan, distribusi kelas ditunjukkan pada Tabel 3.2.

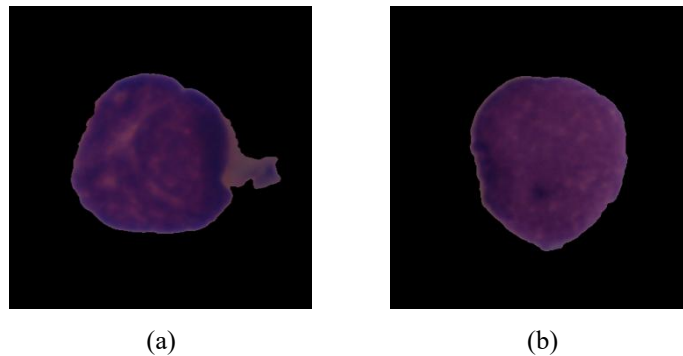
Tabel 3. 2 Distribusi Kelas Dataset

Kelas	Jumlah Citra	Persentase
ALL (Leukemia)	8.491	67,7%
HEM (Normal)	4.037	32,3%
Total	12.528	100%

Berdasarkan Tabel 3.2, terdapat ketidakseimbangan kelas (*class imbalance*) dengan rasio perbandingan sekitar 2:1 antara kelas ALL dan kelas HEM. Kondisi ini perlu diperhatikan dalam proses analisis hasil evaluasi model, khususnya dalam menginterpretasikan nilai metrik *recall* dan *F1-score* pada masing-masing kelas.

Secara visual, perbedaan karakteristik morfologi antara citra kelas ALL dan kelas HEM dapat diamati pada Gambar 3.1. Sel ALL umumnya menampilkan inti sel berukuran besar dengan rasio inti terhadap sitoplasma yang tinggi dan tekstur kromatin yang padat, sedangkan sel HEM menampilkan inti sel yang lebih kecil dengan sitoplasma yang lebih luas dan warna yang lebih terang. Perbedaan

karakteristik visual inilah yang menjadi dasar bagi model dalam melakukan proses ekstraksi fitur dan klasifikasi citra sel darah.



Gambar 3. 1 Citra sel darah (a) Acute Lymphoblastic Leukemia (ALL) dan (b) sel darah normal (HEM)

3.2 Augmentasi Data

Augmentasi data merupakan teknik yang digunakan untuk meningkatkan jumlah dan variasi data latih dengan cara melakukan transformasi tertentu pada citra asli tanpa mengubah label kelasnya. Pada penelitian klasifikasi citra sel darah, augmentasi data berperan penting untuk membantu model dalam mempelajari variasi bentuk, orientasi, dan pola sel darah yang mungkin muncul pada data nyata (Pradana et al., 2024).

Dataset C-NMC 2019 yang dimanfaatkan pada studi ini merupakan dataset publik yang memiliki keterbatasan variasi citra, terutama dari segi orientasi dan posisi sel darah. Kondisi tersebut dapat menyebabkan model deep learning menghadapi overfitting, yaitu model mampu mengenali data latih dengan baik tetapi kurang maksimal saat dihadapkan pada data baru. Oleh karena itu, diperlukan teknik augmentasi data untuk memperkaya karakteristik data latih sehingga model SqueezeNet mampu menjalankan generalisasi dengan lebih baik.

Pada studi ini, augmentasi data diterapkan hanya pada data latih, sementara data uji tidak menghadapi proses augmentasi. Perihal ini bertujuan supaya evaluasi performa model menggambarkan keterampilan model dalam mengklasifikasikan data asli yang belum pernah diamati sebelumnya. Teknik augmentasi yang digunakan meliputi transformasi geometris sederhana yang tetap mempertahankan informasi visual utama dari citra sel darah.

Teknik augmentasi yang diterapkan pada penelitian ini terdiri dari rotasi dan flipping citra. Kedua teknik tersebut dipilih karena sesuai dengan karakteristik citra sel darah, di mana orientasi sel tidak memengaruhi kelasnya (ALL atau HEM). Dengan menerapkan augmentasi ini, diharapkan model SqueezeNet mampu mengenali pola sel darah secara lebih robust terhadap variasi orientasi dan posisi citra.

3.2.1 Rotasi

Augmentasi data menggunakan teknik rotasi dilakukan untuk meningkatkan keberagaman data latih dengan cara memutar citra sel darah pada sudut tertentu terhadap titik pusat citra. Penerapan rotasi bertujuan agar model klasifikasi tidak hanya mengenali pola visual pada orientasi tertentu, tetapi juga mampu mengenali karakteristik sel darah pada berbagai posisi kemiringan. Hal ini penting karena dalam citra mikroskopis, posisi dan orientasi sel darah dapat bervariasi sehingga model CNN perlu memiliki kemampuan mengenali objek pada berbagai orientasi citra (Quiroga et al., 2023).

Secara matematis, transformasi rotasi pada citra dua dimensi dapat direpresentasikan sebagai transformasi koordinat piksel dari titik awal (x,y)

menjadi titik baru (x', y') setelah dilakukan rotasi sebesar sudut θ . Persamaan transformasi rotasi tersebut dinyatakan pada persamaan 3.1.

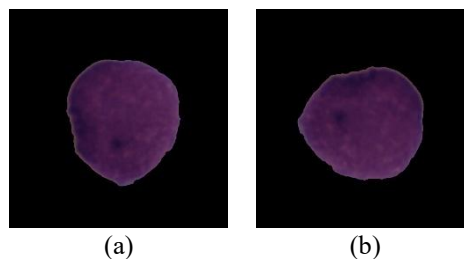
$$x' = x \cos(\theta) - y \sin(\theta) \quad (3.1)$$

$$y' = x \sin(\theta) + y \cos(\theta)$$

Keterangan:

x, y = Koordinat asli piksel
 x', y' = koordinat piksel hasil rotasi
 θ = sudut rotasi

Berdasarkan persamaan tersebut, setiap piksel pada citra asli mengalami perubahan posisi sesuai dengan sudut rotasi yang diberikan. Proses rotasi ini menghasilkan citra baru dengan orientasi yang berbeda, namun tetap mempertahankan karakteristik visual utama dari sel darah yang sama. Contoh hasil rotasi citra ditunjukkan pada Gambar 3.2.



Gambar 3. 2 Hasil Rotasi Citra: (a) citra asli, (b) citra hasil rotasi

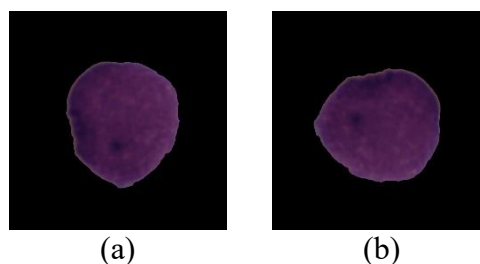
Ilustrasi proses rotasi citra terhadap titik pusat ditunjukkan pada Gambar 3.2, yang menggambarkan perubahan posisi koordinat piksel sebelum dan sesudah dilakukan rotasi. Melalui augmentasi rotasi ini, jumlah variasi data latih meningkat sehingga model SqueezeNet diharapkan mampu melakukan generalisasi yang lebih baik dalam membedakan sel darah penderita Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM).

3.2.2 Flipping

Flipping horizontal merupakan teknik augmentasi data yang dilakukan dengan membalik citra terhadap sumbu vertikal (sumbu Y), sehingga posisi objek pada sisi kiri citra berpindah ke sisi kanan, dan sebaliknya. Teknik ini digunakan untuk meningkatkan variasi data latih agar model klasifikasi mampu mengenali karakteristik sel darah pada berbagai orientasi posisi, mengingat citra mikroskopis sel darah tidak selalu memiliki susunan yang seragam. Secara matematis, transformasi flipping horizontal dapat dinyatakan pada persamaan 3.2.

$$B(x, y) = A(W - x - 1, y) \quad (3.2)$$

Transformasi flipping horizontal menghasilkan citra baru yang merupakan pencerminan dari citra asli terhadap sumbu vertikal. Dengan adanya transformasi ini, pola morfologi sel darah tetap terjaga, namun posisi spasialnya berubah, sehingga model tidak bergantung pada satu orientasi tertentu dalam proses pembelajaran. Hasil penerapan flipping horizontal pada citra sel darah dapat dilihat pada gambar 3.3.



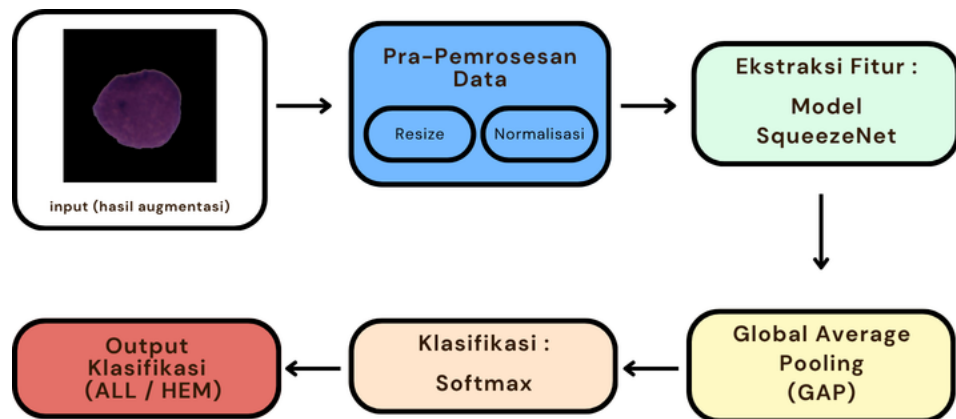
Gambar 3. 3 Hasil flipping horizontal: (a) citra asli, (b) citra hasil flipping horizontal

3.3 Desain Sistem

Desain sistem pada penelitian ini untuk menggambarkan alur kerja sistem klasifikasi citra sel darah putih untuk membedakan antara sel Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM) menggunakan model deep learning SqueezeNet. Sistem yang dibangun difokuskan pada klasifikasi citra, bukan pada proses deteksi objek, sehingga setiap citra yang digunakan telah memiliki label kelas yang jelas dan diproses sebagai satu kesatuan citra utuh.

Alur sistem dimulai dari citra sel darah hasil augmentasi data yang digunakan sebagai data masukan. Citra hasil augmentasi tersebut kemudian diproses pada tahap pra-pemrosesan data untuk menyesuaikan karakteristik citra dengan kebutuhan model, seperti penyeragaman ukuran citra dan normalisasi nilai piksel. Tahap pra-pemrosesan ini bertujuan untuk memastikan bahwa seluruh citra yang digunakan memiliki format yang konsisten sehingga dapat diproses secara optimal oleh model deep learning.

Setelah melalui tahap pra-pemrosesan, citra sel darah selanjutnya dimasukkan ke dalam model SqueezeNet. Pada tahap ini, model melakukan proses ekstraksi fitur secara otomatis melalui lapisan konvolusi dan Fire Module, kemudian dilanjutkan dengan proses klasifikasi untuk menentukan kelas citra. Model SqueezeNet menghasilkan prediksi kelas citra berdasarkan pola fitur yang dipelajari selama proses pelatihan.



Gambar 3. 4 Desain Sistem

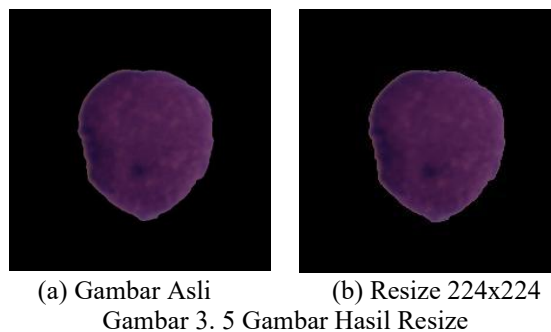
Tahap akhir dari sistem ini adalah keluaran berupa hasil klasifikasi citra sel darah ke dalam salah satu dari dua kelas, yaitu sel ALL atau sel HEM. Hasil keluaran ini digunakan sebagai dasar untuk melakukan evaluasi performa model klasifikasi pada tahap selanjutnya. Secara keseluruhan, alur kerja sistem klasifikasi citra sel darah putih pada penelitian ini dapat dilihat pada Gambar 3.4.

3.3.1 Pra-Pemrosesan Data

Tahap pra-pemrosesan data merupakan tahap awal yang dilakukan sebelum citra sel darah diproses lebih lanjut oleh model SqueezeNet. Pra-pemrosesan tujuannya untuk menyiapkan citra supaya mempunyai format, ukuran, dan karakteristik numerik yang sesuai dengan kebutuhan model, sehingga proses ekstraksi fitur dan klasifikasi mampu berjalan dengan optimal. Pada penelitian ini, tahapan pra-pemrosesan data yang diterapkan mencakup resize citra dan normalisasi nilai piksel.

Proses resize citra dilakukan untuk menyeragamkan ukuran seluruh citra sel darah yang digunakan dalam penelitian. Dataset C-NMC 2019 memiliki citra dengan ukuran awal sebesar 450×450 piksel, sehingga dilakukan proses

pengubahan ukuran menjadi 224×224 piksel agar sesuai dengan ukuran input yang dibutuhkan oleh arsitektur SqueezeNet. Penyesuaian ukuran ini penting karena model Convolutional Neural Network (CNN) umumnya memerlukan dimensi input yang tetap agar dapat memproses data secara konsisten pada setiap lapisan jaringan. Hasil dari resize dapat dilihat pada gambar 3.6.



Selain itu, proses resize juga bertujuan untuk mengurangi kompleksitas komputasi, karena ukuran citra yang lebih kecil akan mempercepat proses pelatihan dan pengujian model tanpa menghilangkan informasi penting secara signifikan. Dengan ukuran citra yang seragam, proses ekstraksi fitur pada setiap lapisan konvolusi dapat dilakukan secara lebih stabil, terstruktur, serta menghasilkan representasi fitur yang lebih konsisten antar citra.

Setelah proses resize, dilakukan normalisasi nilai piksel pada citra. Normalisasi bertujuan untuk mengubah rentang nilai piksel citra ke skala yang lebih kecil dan seragam. Nilai piksel citra digital umumnya berada pada rentang 0 hingga 255, sehingga apabila langsung digunakan sebagai input jaringan saraf, dapat menyebabkan proses pembelajaran menjadi kurang stabil. Oleh karena itu, pada penelitian ini dilakukan normalisasi dengan membagi setiap nilai piksel dengan

nilai maksimum piksel, yaitu 255, sehingga nilai piksel berada pada rentang 0 hingga 1.

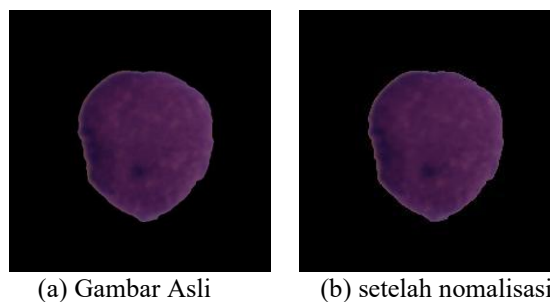
Secara matematis, proses normalisasi nilai piksel dapat dinyatakan pada persamaan 3.3.

$$x' = \frac{x}{255} \quad (3.3)$$

Keterangan:

x = nilai piksel asli citra.
 x' = nilai piksel setelah normalisasi
 255 = nilai maksimum piksel citra

Penerapan normalisasi nilai piksel ini bertujuan untuk membantu mempercepat proses pembelajaran model serta mengurangi perbedaan skala antar fitur yang diekstraksi oleh jaringan saraf konvolusional. Dengan nilai input yang berada pada rentang yang seragam, model SqueezeNet dapat melakukan proses optimasi parameter secara lebih efektif selama tahap pelatihan. Hasil Resize dapat dilihat pada gambar 3.6.



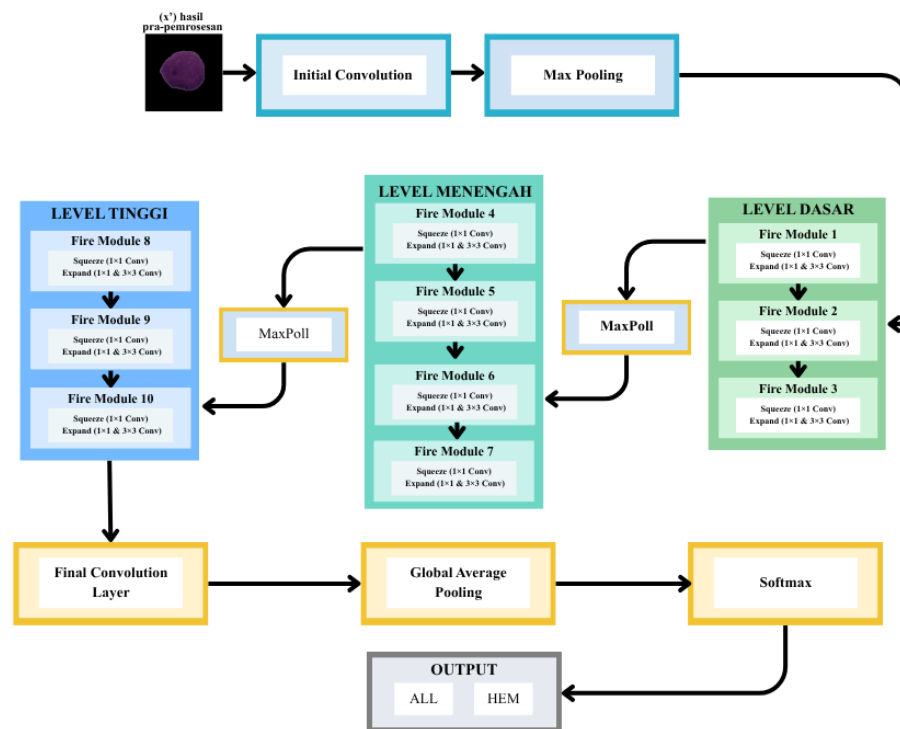
Gambar 3. 6 Normalisasi

Melalui tahapan pra-pemrosesan data yang mencakup resize dan normalisasi ini, citra sel darah yang digunakan sebagai data masukan telah disesuaikan dengan kebutuhan arsitektur SqueezeNet. Tahapan ini diharapkan

dapat mendukung proses klasifikasi citra sel darah leukemia (ALL) dan sel darah normal (HEM) secara lebih akurat dan konsisten.

3.3.2 Arsitektur Model SqueezeNet

Arsitektur model yang digunakan pada penelitian ini adalah SqueezeNet, yaitu jaringan saraf konvolusional yang dirancang untuk menghasilkan performa klasifikasi yang baik dengan jumlah parameter yang relatif kecil (Iandola et al., 2017). Pada penelitian ini, arsitektur SqueezeNet dimodifikasi dengan menambah jumlah *Fire Module* dari 8 menjadi 10, serta mengelompokkannya ke dalam tiga level berdasarkan tingkat abstraksi fitur yang diekstraksi, yaitu Level Dasar (FM 1–3), Level Menengah (FM 4–7), dan Level Tinggi (FM 8–10). Modifikasi ini bertujuan untuk meningkatkan kapasitas ekstraksi fitur model agar lebih optimal dalam membedakan citra sel darah leukemia (ALL) dan sel darah normal (HEM), tanpa mengorbankan efisiensi parameter yang menjadi keunggulan utama SqueezeNet. Arsitektur model SqueezeNet yang digunakan dalam penelitian ini ditunjukkan pada Gambar 3.7.



Gambar 3. 7 Arsitektur Model SqueezeNet

Arsitektur SqueezeNet yang digunakan pada penelitian ini terdiri dari citra *input* hasil pra-pemrosesan, *initial convolution layer*, *max pooling* awal, rangkaian *Fire Module* yang dikelompokkan ke dalam Level Dasar (FM 1–3), Level Menengah (FM 4–7), dan Level Tinggi (FM 8–10), *max pooling* antar level, *final convolution layer*, *global average pooling*, serta lapisan keluaran menggunakan fungsi aktivasi *softmax*. Spesifikasi lengkap setiap *layer* pada arsitektur yang digunakan ditunjukkan pada Tabel 3.3.

Tabel 3. 3 Spesifikasi Arsitektur Model SqueezeNet

Layer	Tipe	Konfigurasi	Output Size
Input	-	$224 \times 224 \times 3$	$224 \times 224 \times 3$
Conv1	Conv2D	64 filter, kernel 3×3 , stride 2, ReLU	$112 \times 112 \times 64$
MaxPool1	MaxPooling2D	Pool 3×3 , stride 2	$55 \times 55 \times 64$
Fire Module 1	Fire Module	Squeeze = 16, Expand = 64	$55 \times 55 \times 128$

Layer	Tipe	Konfigurasi	Output Size
Fire Module 2	Fire Module	Squeeze = 16, Expand = 64	$55 \times 55 \times 128$
Fire Module 3	Fire Module	Squeeze = 32, Expand = 128	$55 \times 55 \times 256$
MaxPool2	MaxPooling2D	Pool 3×3 , stride 2	$27 \times 27 \times 256$
Fire Module 4	Fire Module	Squeeze = 32, Expand = 128	$27 \times 27 \times 256$
Fire Module 5	Fire Module	Squeeze = 48, Expand = 192	$27 \times 27 \times 384$
Fire Module 6	Fire Module	Squeeze = 48, Expand = 192	$27 \times 27 \times 384$
Fire Module 7	Fire Module	Squeeze = 64, Expand = 256	$27 \times 27 \times 512$
MaxPool3	MaxPooling2D	Pool 3×3 , stride 2	$13 \times 13 \times 512$
Fire Module 8	Fire Module	Squeeze = 64, Expand = 256	$13 \times 13 \times 512$
Fire Module 9	Fire Module	Squeeze = 64, Expand = 256	$13 \times 13 \times 512$
Fire Module 10	Fire Module	Squeeze = 64, Expand = 256	$13 \times 13 \times 512$
Conv2 (Final)	Conv2D	2 filter, kernel 1×1	$13 \times 13 \times 2$
GAP	GlobalAveragePooling2D	-	2
Output	Softmax	2 kelas (ALL, HEM)	2

Berdasarkan Gambar 3.7 dan Tabel 3.3 tersebut, setiap komponen pada arsitektur SqueezeNet yang digunakan dalam penelitian ini akan dijelaskan sebagai berikut.

1. Input Citra

Input pada arsitektur SqueezeNet berupa citra sel darah hasil pra-pemrosesan berukuran $224 \times 224 \times 3$ piksel dengan nilai piksel yang telah dinormalisasi ke rentang $[0,1]$. Pada penelitian ini, citra input dinotasikan sebagai x' , yang menunjukkan bahwa citra tersebut bukan citra mentah, melainkan citra yang telah melalui tahapan *resize* dan normalisasi sebagaimana dijelaskan pada Subbab 3.3.1. Citra input ini selanjutnya menjadi masukan awal yang diproses

secara bertahap oleh seluruh lapisan pada arsitektur SqueezeNet, dimulai dari lapisan konvolusi awal hingga lapisan keluaran.

2. Lapisan Konvolusi Awal

Lapisan konvolusi awal merupakan lapisan pertama yang memproses citra input pada arsitektur SqueezeNet. Lapisan ini memanfaatkan 64 filter berukuran 3×3 dengan stride 2 dan fungsi aktivasi ReLU, sehingga menghasilkan *feature map* berukuran $112 \times 112 \times 64$. Lapisan ini berfungsi untuk mengekstraksi fitur dasar dari citra input, seperti tepi sel, tekstur permukaan, dan pola intensitas awal pada citra sel darah. Secara matematis, operasi konvolusi pada lapisan ini mampu diamati pada persamaan 3.4.

$$y(i, j) = \sum_m \sum_n x'(i + m, j + n) \cdot w(m, n) \quad (3.4)$$

Keterangan:

$x'(i + m, j + n)$	=	nilai piksel citra hasil pra-pemrosesan
$w(m, n)$	=	bobot atau kernel konvolusi
$y(i, j)$	=	nilai keluaran hasil konvolusi
m dan n	=	indeks pergeseran kernel

Melalui operasi konvolusi tersebut, lapisan ini mengekstraksi informasi fitur dari citra input secara bertahap pada seluruh area citra.

3. Max Pooling Awal

Setelah lapisan konvolusi awal, *feature map* berukuran $112 \times 112 \times 64$ selanjutnya diproses oleh lapisan Max Pooling pertama. Lapisan ini menggunakan pool size 3×3 dengan stride 2, sehingga mereduksi ukuran *feature map* menjadi $55 \times 55 \times 64$. Tujuan dari lapisan ini adalah untuk mempertahankan fitur yang paling dominan dari setiap area *feature map*

sekaligus mengurangi dimensi spasial, sehingga kompleksitas komputasi pada lapisan selanjutnya dapat ditekan.

4. Fire Module

Fire Module ialah komponen utama dalam arsitektur SqueezeNet yang dibuat untuk mengurangi jumlah parameter jaringan tanpa mengurangi kemampuan representasi fitur. Setiap Fire Module terbagi dari dua bagian utama, yakni *squeeze layer* dan *expand layer* (Iandola et al., 2017), sebagaimana ditunjukkan pada Gambar 3.7.

a. Squeeze Layer

Squeeze layer merupakan bagian pertama dari *Fire Module* yang menggunakan operasi konvolusi berukuran 1×1 . Tujuan dari lapisan ini adalah untuk mengurangi jumlah *channel* pada *feature map* hasil lapisan sebelumnya, sehingga beban komputasi pada lapisan *expand* berikutnya dapat ditekan. Proses reduksi *channel* ini dilakukan tanpa menghilangkan informasi fitur penting secara signifikan.

Operasi pada *squeeze layer* dapat dirumuskan pada persamaan 3.5.

$$S = f(x' * w_s) \quad (3.5)$$

Keterangan:

w_s	=	kernel konvolusi berukuran 1×1
$f(\cdot)$	=	fungsi aktivasi ReLU
S	=	feature map hasil <i>squeeze layer</i>

Melalui operasi ini, *squeeze layer* mereduksi jumlah *channel* pada *feature map* sehingga beban komputasi pada lapisan *expand* berikutnya menjadi lebih ringan tanpa menghilangkan informasi fitur penting secara signifikan.

b. Expand Layer

Expand layer merupakan bagian kedua dari *Fire Module* yang berfungsi untuk memperkaya representasi fitur. Lapisan ini menerapkan dua jenis konvolusi secara paralel, yaitu konvolusi 1×1 dan konvolusi 3×3 . Konvolusi 1×1 bertugas menangkap hubungan antar *channel*, sedangkan konvolusi 3×3 bertugas menangkap pola spasial pada *feature map*. Hasil dari kedua operasi tersebut kemudian digabungkan (*concatenation*) untuk membentuk *feature map* yang lebih informatif dan kaya fitur. Secara matematis, operasi pada *expand layer* dapat dirumuskan pada persamaan 3.6.

$$E = f(S * w_{e1}), f(S * w_{e3}) \quad (3.6)$$

Keterangan:

w_{e1}	=	kernel konvolusi 1×1
w_{e3}	=	kernel konvolusi 3×3
E	=	feature map hasil <i>expand layer</i>

Melalui operasi ini, *expand layer* memperkaya representasi fitur dengan menggabungkan hasil ekstraksi dari dua jenis konvolusi yang berbeda secara paralel.

c. Fungsi dan Perbedaan Antar Fire Module

Pada arsitektur SqueezeNet yang digunakan dalam penelitian ini, terdapat 10 *Fire Module* yang dikelompokkan ke dalam tiga level berdasarkan tingkat abstraksi fitur yang diekstraksi. Penambahan jumlah *Fire Module* dari 8 (standar) menjadi 10 bertujuan untuk memperdalam jaringan sehingga model memiliki kapasitas ekstraksi fitur yang lebih kaya, tanpa menambah parameter secara signifikan karena setiap *Fire Module* tetap menggunakan mekanisme *squeeze layer* dan *expand layer*. Perbedaan antar *Fire Module* terletak pada jumlah filter yang digunakan, yang meningkat secara bertahap seiring

bertambahnya kedalaman jaringan untuk memungkinkan ekstraksi fitur yang makin kompleks dan abstrak.

1) Level Dasar (Fire Module 1–3)

Level Dasar terdiri dari Fire Module 1 hingga Fire Module 3 dengan konfigurasi squeeze = 16-32 dan expand = 64-128, menghasilkan feature map berukuran 55×55 . Level ini berfungsi untuk mengekstraksi fitur tingkat rendah dari citra sel darah, seperti tepi sel, kontur inti sel, dan perbedaan intensitas dasar antara inti dan sitoplasma. Fitur-fitur ini menjadi dasar bagi proses ekstraksi fitur pada lapisan selanjutnya.

Setelah Fire Module 3, feature map berukuran $55 \times 55 \times 256$ diproses oleh lapisan Max Pooling kedua menggunakan pool size 3×3 dengan stride 2, sehingga ukuran feature map tereduksi menjadi $27 \times 27 \times 256$. Proses ini berfungsi untuk mereduksi dimensi spasial sebelum memasuki Level Menengah, sehingga model dapat memfokuskan ekstraksi fitur pada pola yang lebih kompleks dengan beban komputasi yang lebih ringan.

2) Level Menengah (Fire Module 4–7)

Level Menengah terdiri dari Fire Module 4 hingga Fire Module 7 dengan konfigurasi squeeze=32–64 dan expand=128–256, menghasilkan feature map berukuran 27×27 . Level ini berperan dalam mengekstraksi fitur tingkat menengah dengan mengombinasikan fitur dasar yang telah diperoleh sebelumnya. Pada tahap ini, model mulai menangkap variasi bentuk inti sel, tekstur kromatin, serta pola morfologi yang lebih

kompleks, yang penting untuk membedakan sel leukemia (ALL) dan sel darah normal (HEM).

Setelah Fire Module 7, feature map berukuran $27 \times 27 \times 512$ diproses oleh lapisan Max Pooling ketiga menggunakan pool size 3×3 dengan stride 2, sehingga ukuran feature map tereduksi menjadi $13 \times 13 \times 512$. Reduksi dimensi ini mempersiapkan feature map untuk memasuki Level Tinggi, di mana model akan mengekstraksi fitur yang lebih abstrak dan spesifik untuk mendukung proses klasifikasi akhir.

3) Level Tinggi (Fire Module 8–10)

Level Tinggi terdiri dari Fire Module 8 hingga Fire Module 10 dengan konfigurasi $\text{squeeze}=64$ dan $\text{expand}=256$, menghasilkan feature map berukuran 13×13 . Level ini berfungsi untuk mengekstraksi fitur tingkat tinggi yang bersifat lebih abstrak dan spesifik terhadap proses klasifikasi. Pada lapisan ini, model mengintegrasikan seluruh informasi fitur sebelumnya untuk membentuk representasi global citra sel darah yang mendukung penentuan kelas ALL dan HEM secara lebih akurat.

Secara keseluruhan, penyusunan Fire Module secara bertahap memungkinkan SqueezeNet melakukan ekstraksi fitur secara progresif, dari fitur sederhana hingga fitur kompleks, sehingga meningkatkan kemampuan model dalam membedakan kedua kelas citra sel darah.

5. Final Convolution

Setelah seluruh rangkaian *Fire Module* selesai mengekstraksi fitur, *feature map* yang dihasilkan kemudian diproses oleh *final convolution layer*.

Lapisan ini memanfaatkan 2 filter berukuran 1×1 untuk menyesuaikan jumlah *channel* dengan total kelas yang digunakan pada studi ini, yaitu ALL dan HEM, sehingga menghasilkan *feature map* berukuran $13 \times 13 \times 2$. Penggunaan konvolusi 1×1 pada tahap ini memungkinkan penyesuaian dimensi *channel* tanpa mengubah dimensi spasial *feature map*.

Secara matematis, operasi pada *final convolution layer* dapat dirumuskan pada persamaan 3.7.

$$F = f(E * w_f) \quad (3.7)$$

Keterangan :

- E = Feature map keluaran dari Fire Module terakhir
- w_f = Kernel konvolusi berukuran 1×1
- $f(\cdot)$ = Fungsi aktivasi ReLU
- F = Feature map hasil final convolution layer berukuran $13 \times 13 \times 2$

Melalui operasi ini, *final convolution layer* mereduksi jumlah *channel feature map* menjadi 2, yang masing-masing merepresentasikan kelas ALL dan HEM, tanpa mengubah dimensi spasial *feature map*.

6. Global Average Pooling (GAP)

Feature map berukuran $13 \times 13 \times 2$ yang dihasilkan dari *final convolution layer* selanjutnya diproses oleh lapisan *Global Average Pooling* (GAP). Lapisan ini bekerja dengan menghitung nilai rata-rata seluruh elemen pada setiap *channel feature map*, sehingga menghasilkan vektor berukuran 2 yang merepresentasikan informasi global dari masing-masing kelas. GAP menggantikan *fully connected layer* yang biasa digunakan pada arsitektur CNN konvensional, sehingga jumlah parameter model berkurang secara signifikan dan risiko *overfitting* dapat ditekan.

Secara matematis, operasi Global Average Pooling dapat dirumuskan pada persamaan 3.8.

$$G_k = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W F_k(i, j) \quad (3.8)$$

Keterangan:

$F_k(i, j)$	=	feature map ke- k
H dan W	=	tinggi dan lebar feature map
G_k	=	nilai rata-rata feature map ke- k

Melalui operasi ini, GAP menghasilkan representasi global dari setiap *channel feature map* yang berikutnya dimanfaatkan sebagai input pada lapisan klasifikasi, sekaligus membantu meminimalisir risiko *overfitting* dan menumbuhkan keterampilan generalisasi model.

7. Softmax dan Output Klasifikasi

Lapisan keluaran pada arsitektur SqueezeNet ini menggunakan fungsi aktivasi *softmax* yang menerima vektor berukuran 2 hasil dari GAP secara langsung. Fungsi *softmax* mengubah nilai tersebut menjadi probabilitas untuk tiap-tiap kelas, yaitu ALL dan HEM, dengan total probabilitas seluruh kelas bernilai 1. Fungsi softmax dirumuskan pada persamaan 3.9.

$$\text{Softmax}(z_k) = \frac{e^{z_k}}{\sum_{j=1}^C e^{z_j}} \quad (3.9)$$

Keterangan:

z_k	=	nilai keluaran neuron ke- k sebelum softmax
C	=	jumlah kelas
e	=	bilangan eksponensial

Kelas dengan nilai probabilitas tertinggi kemudian ditetapkan sebagai hasil klasifikasi akhir dari citra sel darah yang diinputkan.

3.4 Evaluasi Kinerja Model

Evaluasi kinerja model dijalankan untuk menilai keterampilan arsitektur SqueezeNet dalam mengklasifikasikan citra sel darah ke dalam dua kelas, yakni Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM). Evaluasi ini dilakukan berdasarkan hasil prediksi model kepada data uji, sehingga performa model dapat dinilai dengan cara objektif. Proses evaluasi kinerja model pada studi ini dijalankan melewati sejumlah tahapan dibawah ini:

a) Penyusunan Confusion Matrix

Tahap awal evaluasi dilakukan dengan menyusun confusion matrix, yaitu matriks yang digunakan untuk membandingkan hasil prediksi model dengan label sebenarnya pada data uji. Confusion matrix digunakan untuk mengetahui jumlah prediksi yang benar dan salah pada masing-masing kelas. Pada penelitian klasifikasi dua kelas (ALL dan HEM), confusion matrix terdiri dari empat komponen utama yang ditunjukkan pada Tabel 3.4.

Tabel 3. 4 Confusion Matrix Klasifikasi Citra Sel Darah

Actual Class		Assigned Class	
		ALL	HEM
	ALL	TP	FN
	HEM	FP	TN

Keterangan:

True Positive (TP) = citra sel ALL yang diprediksi benar sebagai ALL
 True Negative (TN) = citra sel HEM yang diprediksi benar sebagai HEM
 False Positive (FP) = citra sel HEM yang salah diprediksi sebagai ALL
 False Negative (FN) = citra sel ALL yang salah diprediksi sebagai HEM

b) Perhitungan Akurasi

Akurasi digunakan untuk mengukur tingkat ketepatan model dalam mengklasifikasikan seluruh data uji secara keseluruhan. Nilai akurasi diperoleh

dari perbandingan jumlah prediksi yang benar terhadap total data uji. Akurasi dirumuskan pada persamaan 3.10.

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.10)$$

c) Perhitungan Precision

Precision dimanfaatkan untuk menilai tingkat ketepatan model dalam memperkirakan suatu kelas tertentu. Pada penelitian ini, precision menunjukkan seberapa tepat model dalam memprediksi citra sebagai kelas ALL atau HEM, berdasarkan hasil prediksi yang dihasilkan. Perhitungan nilai precision dirumuskan pada persamaan 3.11.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.11)$$

d) Perhitungan Recall

Recall digunakan untuk mengukur kemampuan model dalam mengenali seluruh data yang termasuk ke dalam suatu kelas tertentu. Dalam konteks penelitian ini, recall menunjukkan kemampuan model dalam mendeteksi citra sel darah leukemia (ALL) atau sel darah normal (HEM) secara menyeluruh. Perhitungan nilai recall dirumuskan pada persamaan 3.12.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.12)$$

e) Perhitungan F1-Score

F1-score dimanfaatkan sebagai metrik evaluasi tambahan untuk memberi penilaian yang lebih seimbang antara precision dan recall. Metrik ini sangat berguna dalam klasifikasi dua kelas, terutama saat distribusi data antar kelas

tidak seimbang. Perhitungan nilai F1-score dirumuskan pada persamaan 3.13.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.13)$$

Melalui tahapan evaluasi tersebut, kinerja model SqueezeNet dalam mengklasifikasikan citra sel darah leukemia dan sel darah normal dapat dinilai secara terukur dan sistematis.

BAB IV

HASIL DAN PEMBAHASAN

4.1 Skenario Uji Coba

Pada penelitian ini, dilakukan pengujian model SqueezeNet untuk mengklasifikasikan citra sel darah Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM) menggunakan dataset C-NMC 2019 dengan total 12.528 citra. Dataset dibagi menjadi data pelatihan dan data pengujian menggunakan tiga variasi rasio pembagian data, yaitu 70:30, 80:20, dan 90:10. Variasi rasio tersebut digunakan untuk menganalisis pengaruh jumlah data pelatihan terhadap kemampuan model dalam menghasilkan prediksi yang akurat. Pembagian data pada masing-masing skenario ditunjukkan pada Tabel 4.1.

Tabel 4. 1 Skenario Pembagian Data

Rasio	Jumlah Data				Total Data
	Training		Testing		
	ALL	HEM	ALL	HEM	
70:30	5.943	2.825	2.548	1.212	12.528
80:20	6.792	3.229	1.699	808	12.528
90:10	7.641	3.633	850	404	12.528

Setelah pembagian data dilakukan, proses augmentasi diterapkan pada data pelatihan menggunakan teknik rotasi dan *horizontal flipping*. Augmentasi bertujuan untuk meningkatkan variasi data latih sehingga model dapat mempelajari karakteristik citra secara lebih baik dan memiliki kemampuan generalisasi yang lebih baik terhadap data yang belum pernah dilihat sebelumnya. Data pengujian tidak mengalami proses augmentasi agar evaluasi performa model dilakukan

menggunakan data yang merepresentasikan kondisi sebenarnya, sehingga tidak terjadi *data leakage* yang dapat memengaruhi keandalan hasil evaluasi. Hasil augmentasi data latih pada setiap skenario ditunjukkan pada Tabel 4.2.

Tabel 4. 2 Jumlah Data Latih Kelas HEM Setelah Augmentasi

Rasio Data	Jumlah Data	
	Sebelum Augmentasi	Setelah Augmentasi
70:30	2.825	5.943
80:20	3.229	6.792
90:10	3.633	7.641

Seluruh citra yang digunakan dalam proses pelatihan dan pengujian menjalani tahap pra-pemrosesan berupa *resize* citra menjadi 224×224 piksel dan normalisasi nilai piksel ke rentang [0,1] sesuai dengan kebutuhan input arsitektur SqueezeNet.

Pada setiap skenario pembagian data, dilakukan variasi parameter pelatihan berupa *batch size* sebesar 16, 32, dan 64 untuk menganalisis pengaruh ukuran *batch* terhadap performa model. Model dilatih menggunakan optimizer Adam dengan *learning rate* sebesar 0,0001 selama 50 *epoch* menggunakan Google Colab dengan dukungan GPU Tesla T4. Seluruh kombinasi skenario pengujian yang dapat dilihat pada Tabel 4.3.

Tabel 4. 3 Skenario Uji Coba

Skenario	Rasio Data	Batch Size	Epoch
1	70:30	16,32,64	50
2	80:20	16,32,64	50
3	90:10	16,32,64	50

Evaluasi performa model dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* untuk menilai kemampuan model dalam mengklasifikasikan citra sel darah ALL dan HEM pada setiap skenario pengujian.

Selain itu, model SqueezeNet yang digunakan memiliki total 1.117.896 parameter dengan ukuran model sebesar 4,26 MB, yang menunjukkan bahwa modifikasi arsitektur yang dilakukan tetap mempertahankan karakteristik efisiensi parameter sebagai keunggulan utama SqueezeNet.

4.2 Hasil Uji Coba

Pada penelitian ini, dilakukan uji coba model SqueezeNet untuk mengklasifikasikan citra sel darah Acute Lymphoblastic Leukemia (ALL) dan sel darah normal (HEM) menggunakan variasi rasio pembagian data dan batch size.

Setiap skenario pengujian menggunakan tiga variasi batch size, yaitu 16, 32, dan 64. Evaluasi performa model dilakukan menggunakan metrik accuracy, precision, recall, dan F1-score untuk mengetahui kemampuan model dalam melakukan klasifikasi.

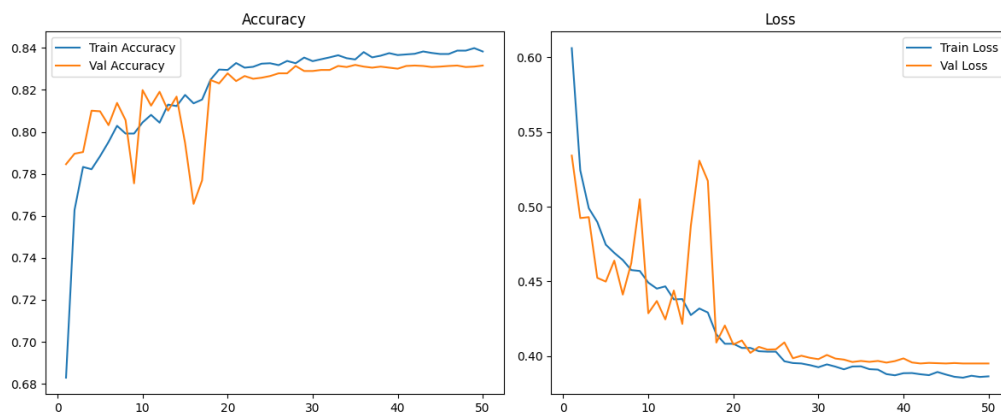
Hasil pengujian pada setiap skenario disajikan dalam bentuk tabel dan grafik untuk menunjukkan performa model berdasarkan variasi rasio pembagian data dan batch size yang digunakan.

4.2.1 Hasil Uji Coba Skenario 1

Skenario 1 menggunakan pembagian data dengan rasio 70:30, di mana 70% data digunakan sebagai data latih dan 30% sebagai data uji. Dari total 12.528 citra, sebanyak 8.768 citra digunakan untuk pelatihan dan 3.760 citra digunakan untuk pengujian. Pada skenario ini, model dilatih selama 50 epoch dengan tiga variasi *batch size*, yaitu 16, 32, dan 64, untuk mengetahui pengaruh *batch size* terhadap performa model.

a. Batch Size 16

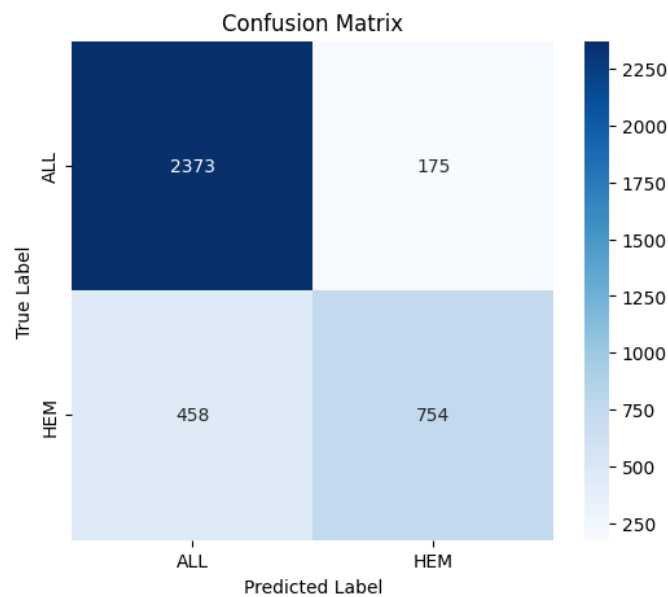
Pada pengujian dengan *batch size* 16, model dilatih selama 50 epoch menggunakan 8.768 citra latih dan diuji pada 3.760 citra uji. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.1.



Gambar 4. 1 Grafik Pelatihan Model Skenario 1 Batch Size 16

Berdasarkan Gambar 4.1, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 68,29% dan *validation accuracy* dari 78,46%, kemudian keduanya meningkat secara bertahap hingga mencapai *training accuracy* sebesar 83,87% dan *validation accuracy* sebesar 83,16% pada epoch ke-50. Kedua kurva relatif berdekatan sepanjang proses pelatihan, mengindikasikan bahwa model tidak mengalami *overfitting* secara signifikan. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,6061 pada epoch pertama menjadi 0,3865 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5342 menjadi 0,3951. Kedua kurva menunjukkan pola konvergensi yang

stabil pada epoch-epoch akhir, mengindikasikan bahwa model telah berhasil mempelajari pola fitur citra sel darah dengan baik. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.2.



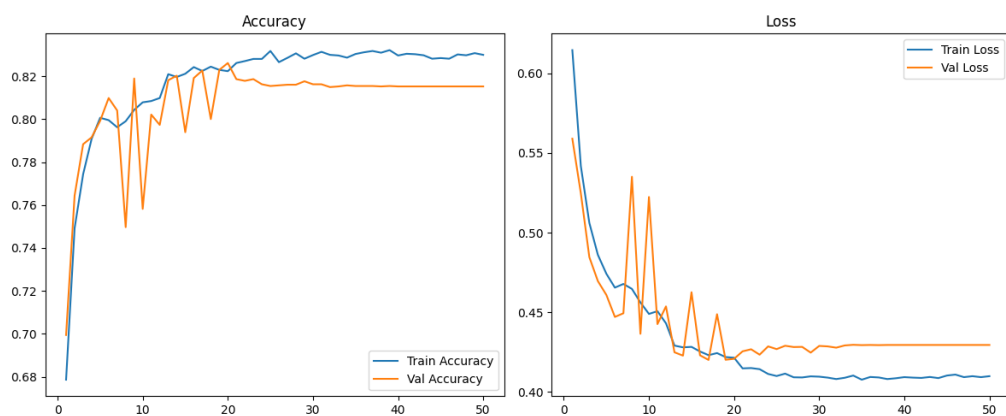
Gambar 4. 2 Confusion Matrix Skenario 1 Batch Size 16

Berdasarkan Gambar 4.2, dari 2.548 citra ALL, sebanyak 2.373 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 175 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 1.212 citra HEM, sebanyak 754 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 458 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 16 menghasilkan akurasi sebesar 83,16% dengan *precision* kelas ALL sebesar 0,84, *recall* sebesar 0,93, dan *F1-score* sebesar 0,88. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,81, *recall* sebesar 0,62, dan *F1-score* sebesar 0,70.

Tingginya *recall* kelas ALL (0,93) menunjukkan bahwa model mampu mendeteksi sebagian besar sel leukemia dengan baik. Namun, nilai *recall* kelas HEM yang lebih rendah (0,62) mengindikasikan bahwa model masih cukup sering salah mengklasifikasikan sel darah normal sebagai ALL. Hal ini dapat disebabkan oleh ketidakseimbangan jumlah data antara kelas ALL (8.491 citra) dan kelas HEM (4.037 citra), sehingga model cenderung lebih bias terhadap kelas mayoritas.

b. Batch Size 32

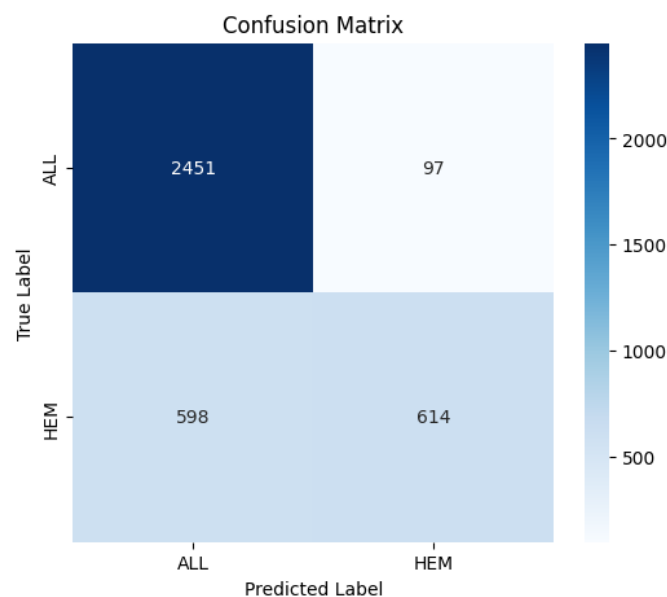
Pada pengujian dengan *batch size* 32, model dilatih selama 50 epoch menggunakan 8.768 citra latih dan diuji pada 3.760 citra uji. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.3.



Gambar 4. 3 Grafik Pelatihan Model Skenario 1 Batch Size 32

Berdasarkan Gambar 4.3, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 67,87% dan *validation accuracy* dari 69,95%, kemudian keduanya meningkat secara bertahap hingga

mencapai *training accuracy* sebesar 82,99% dan *validation accuracy* sebesar 81,52% pada epoch ke-50. Kedua kurva relatif berdekatan sepanjang proses pelatihan, mengindikasikan bahwa model tidak mengalami *overfitting* secara signifikan. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,6146 pada epoch pertama menjadi 0,4099 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5590 menjadi 0,4295. Kedua kurva menunjukkan pola konvergensi yang stabil pada epoch-epoch akhir, mengindikasikan bahwa model telah berhasil mempelajari pola fitur citra sel darah dengan baik. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.4.



Gambar 4. 4 Confusion Matrix Skenario 1 Batch Size 32

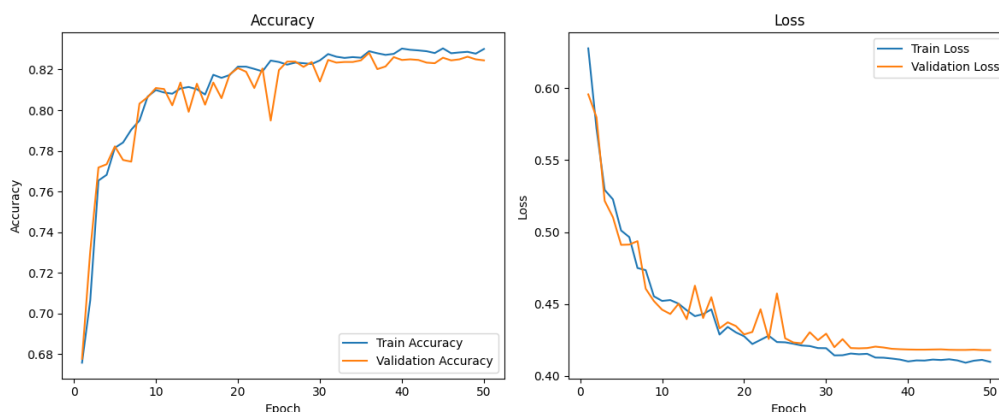
Berdasarkan Gambar 4.4, dari 2.548 citra ALL, sebanyak 2.451 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 97 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 1.212 citra HEM,

sebanyak 614 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 598 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 32 menghasilkan akurasi sebesar 81,52% dengan *precision* kelas ALL sebesar 0,80, *recall* sebesar 0,96, dan *F1-score* sebesar 0,88. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,86, *recall* sebesar 0,51, dan *F1-score* sebesar 0,64.

Batch size 32 menghasilkan *recall* ALL tertinggi di antara ketiga variasi (0,96), namun dengan konsekuensi *recall* HEM yang paling rendah (0,51). Hal ini menunjukkan bahwa model dengan *batch size* 32 sangat sensitif dalam mendeteksi sel leukemia, tetapi kurang optimal dalam mengenali sel darah normal. *Precision* HEM yang tinggi (0,86) mengindikasikan bahwa ketika model memprediksi HEM, prediksi tersebut cenderung benar, namun jumlah prediksi HEM yang dihasilkan sangat sedikit. Sama seperti *batch size* 16, ketidakseimbangan jumlah data antara kelas ALL dan HEM diduga menjadi faktor utama yang memengaruhi performa model pada kelas HEM.

c. Batch Size 64

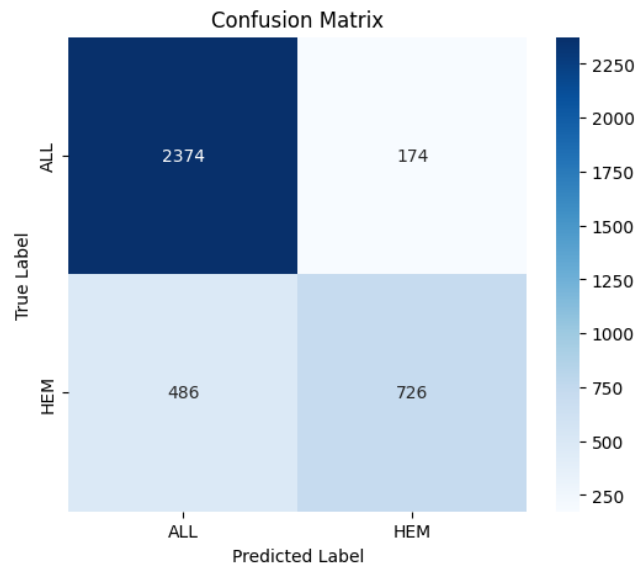
Pada pengujian dengan *batch size* 64, model dilatih selama 50 epoch menggunakan 8.768 citra latih dan diuji pada 3.760 citra uji. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.5.



Gambar 4. 5 Grafik Pelatihan Model Skenario 1 Batch Size 64

Berdasarkan Gambar 4.5, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 67,58% dan *validation accuracy* dari 67,77%, kemudian keduanya meningkat secara bertahap hingga mencapai *training accuracy* sebesar 83,01% dan *validation accuracy* sebesar 82,45% pada epoch ke-50. Dibandingkan *batch size* 16 dan 32, kurva *batch size* 64 menunjukkan peningkatan yang lebih lambat pada epoch-epoch awal, hal ini wajar karena *batch size* yang lebih besar menyebabkan pembaruan bobot yang lebih jarang per epoch sehingga model membutuhkan lebih banyak iterasi untuk konvergen. Kedua kurva relatif berdekatan sepanjang proses pelatihan, mengindikasikan bahwa model tidak mengalami *overfitting* secara signifikan. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,6278 pada epoch pertama menjadi 0,4097 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5958 menjadi 0,4179. Kedua kurva menunjukkan pola konvergensi yang stabil pada epoch-epoch akhir, mengindikasikan bahwa model telah berhasil mempelajari pola

fitur citra sel darah dengan baik. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.6.



Gambar 4. 6 Confusion Matrix Skenario 1 Batch Size 64

Berdasarkan Gambar 4.6, dari 2.548 citra ALL, sebanyak 2.374 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 174 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 1.212 citra HEM, sebanyak 726 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 486 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 64 menghasilkan akurasi sebesar 82,45% dengan *precision* kelas ALL sebesar 0,83, *recall* sebesar 0,93, dan *F1-score* sebesar 0,88. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,81, *recall* sebesar 0,60, dan *F1-score* sebesar 0,69.

Pola performa *batch size* 64 cukup serupa dengan *batch size* 16, di mana *recall* ALL tinggi (0,93) namun *recall* HEM masih belum optimal (0,60). Meskipun demikian, *batch size* 64 menghasilkan keseimbangan yang sedikit lebih baik

antara kelas ALL dan HEM dibandingkan *batch size* 32, yang terlihat dari nilai *recall* HEM yang lebih tinggi (0,60 vs 0,51). Ketidakseimbangan jumlah data antara kelas ALL dan HEM tetap menjadi faktor dominan yang memengaruhi performa model pada ketiga variasi *batch size*.

d. Perbandingan Hasil Skenario 1

Setelah dilakukan serangkaian pengujian dengan tiga variasi *batch size*, yaitu *batch size* 16, 32, dan 64 pada Skenario 1 dengan rasio pembagian data 70:30, seluruh hasil evaluasi yang meliputi akurasi, *precision*, *recall*, dan *F1-score* dari masing-masing variasi *batch size* dirangkum secara keseluruhan pada Tabel 4.4

Tabel 4. 4 Perbandingan Hasil Evaluasi Skenario 1 (70:30)

Batch Size	Accuracy	Precision	Recall	F1-Score
16	83,16%	82,5%	77,5%	79%
32	81,52%	83%	73,5%	76%
64	82,45%	82%	76,5%	78,5%

Berdasarkan Tabel 4.4, *batch size* 16 menghasilkan performa terbaik pada Skenario 1 dengan akurasi 83,16%, *precision* 82,5%, *recall* 77,5%, dan *F1-score* 79%. Meskipun *batch size* 32 menghasilkan *precision* tertinggi (83%), nilai *recall* yang paling rendah (73,5%) menunjukkan bahwa model terlalu banyak melewati prediksi yang benar, khususnya pada kelas HEM. *Batch size* 64 menghasilkan performa yang cukup kompetitif dengan *F1-score* 78,5%, namun masih di bawah *batch size* 16 pada seluruh metrik evaluasi.

Secara keseluruhan, hasil Skenario 1 menunjukkan bahwa model SqueezeNet mampu mengklasifikasikan citra sel darah dengan akurasi yang cukup baik pada rasio pembagian data 70:30. Namun, ketidakseimbangan performa antara

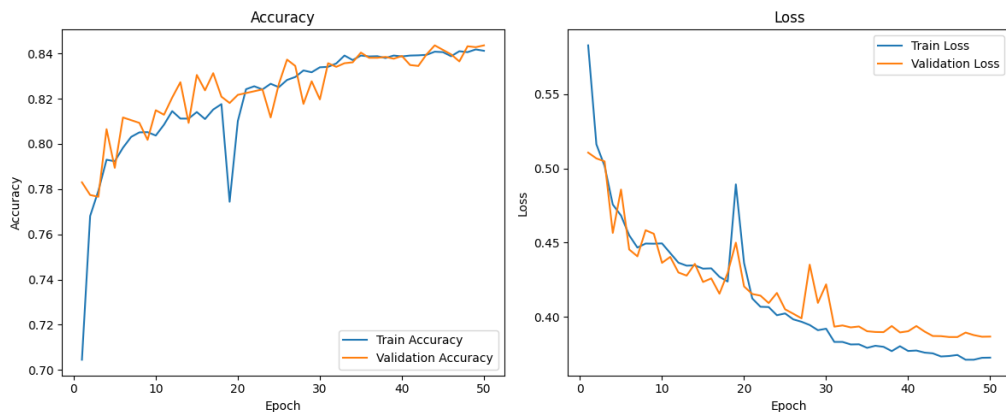
kelas ALL dan HEM pada seluruh variasi batch size mengindikasikan bahwa jumlah data latih yang lebih besar kemungkinan dapat meningkatkan kemampuan generalisasi model, khususnya untuk kelas HEM yang memiliki jumlah data lebih sedikit. Hal ini akan dibuktikan pada hasil pengujian Skenario 2 dan Skenario 3 dengan rasio pembagian data yang lebih besar.

4.2.2 Hasil Uji Coba Skenario 2

Skenario 2 menggunakan pembagian data dengan rasio 80:20, di mana 80% data digunakan sebagai data latih dan 20% sebagai data uji. Dari total 12.528 citra, sebanyak 10.021 citra digunakan untuk pelatihan dan 2.507 citra digunakan untuk pengujian. Pada skenario ini, model dilatih selama 50 epoch dengan tiga variasi *batch size*, yaitu 16, 32, dan 64, untuk mengetahui pengaruh *batch size* terhadap performa model.

a. Batch Size 16

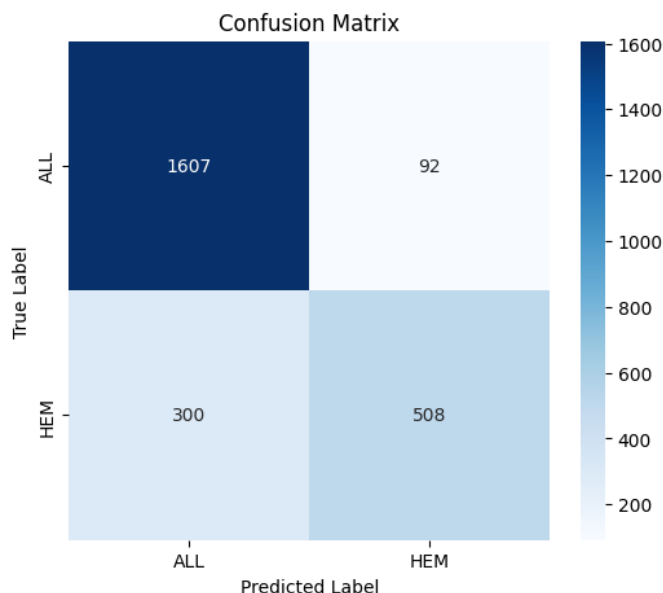
Pada pengujian dengan *batch size* 16, model dilatih selama 50 epoch menggunakan 10.021 citra latih dan diuji pada 2.507 citra uji. Jumlah data latih pada skenario ini lebih besar dibandingkan Skenario 1, yaitu bertambah sebanyak 1.253 citra, sehingga diharapkan model dapat mempelajari pola fitur citra sel darah yang lebih beragam dan menghasilkan kemampuan generalisasi yang lebih baik. Dengan *batch size* 16 dan jumlah data latih sebesar 10.021 citra, model melakukan sebanyak 627 iterasi pembaruan bobot per epoch selama proses pelatihan berlangsung. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.7.



Gambar 4. 7 Grafik Pelatihan Model Skenario 2 Batch Size 16

Berdasarkan Gambar 4.7, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 70,45% dan *validation accuracy* dari 78,30%, kemudian keduanya meningkat secara bertahap hingga mencapai *training accuracy* sebesar 84,12% dan *validation accuracy* sebesar 84,36% pada epoch ke-50. Terdapat fluktuasi yang cukup mencolok pada sekitar epoch ke-18 hingga epoch ke-19, di mana *training accuracy* mengalami penurunan sementara sebelum kembali meningkat dan akhirnya stabil pada epoch-epoch berikutnya. Setelah epoch ke-20, kedua kurva mulai berdekatan dan konvergen secara stabil, mengindikasikan bahwa model tidak mengalami *overfitting* secara signifikan. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,5830 pada epoch pertama menjadi 0,3725 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5107 menjadi 0,3867. Kedua kurva menunjukkan pola konvergensi yang stabil pada epoch-epoch akhir, mengindikasikan bahwa model telah berhasil mempelajari pola fitur citra sel darah dengan baik.

Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.8.



Gambar 4. 8 Confusion Matrix Skenario 2 Batch Size 16

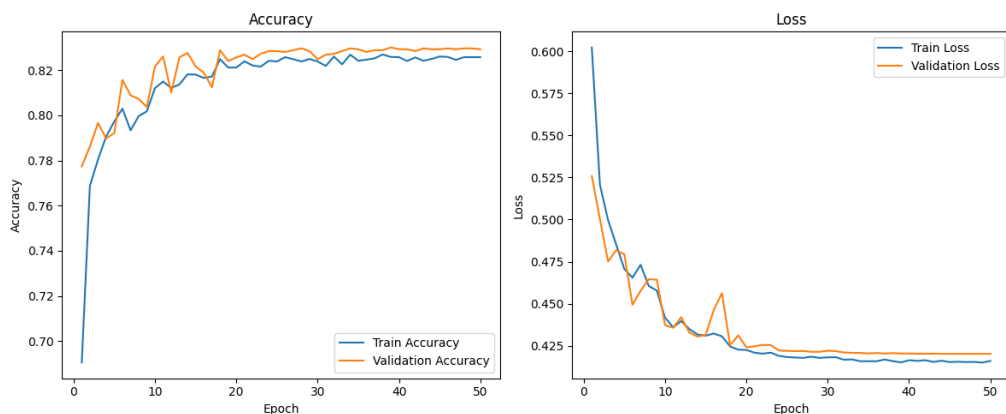
Berdasarkan Gambar 4.8, dari 1.699 citra ALL, sebanyak 1.607 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 92 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 808 citra HEM, sebanyak 508 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 300 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 16 menghasilkan akurasi sebesar 84,36% dengan *precision* kelas ALL sebesar 0,84, *recall* sebesar 0,95, dan *F1-score* sebesar 0,89. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,85, *recall* sebesar 0,63, dan *F1-score* sebesar 0,72.

Dibandingkan dengan Skenario 1, hasil *batch size* 16 pada Skenario 2 menunjukkan peningkatan performa yang cukup signifikan, di mana akurasi meningkat dari 83,16% menjadi 84,36% dan *F1-score* kelas HEM meningkat

dari 0,70 menjadi 0,72. Hal ini mengindikasikan bahwa penambahan jumlah data latih dari 70% menjadi 80% memberikan dampak positif terhadap kemampuan generalisasi model, khususnya dalam mengklasifikasikan kelas HEM yang memiliki jumlah data lebih sedikit.

b. Batch Size 32

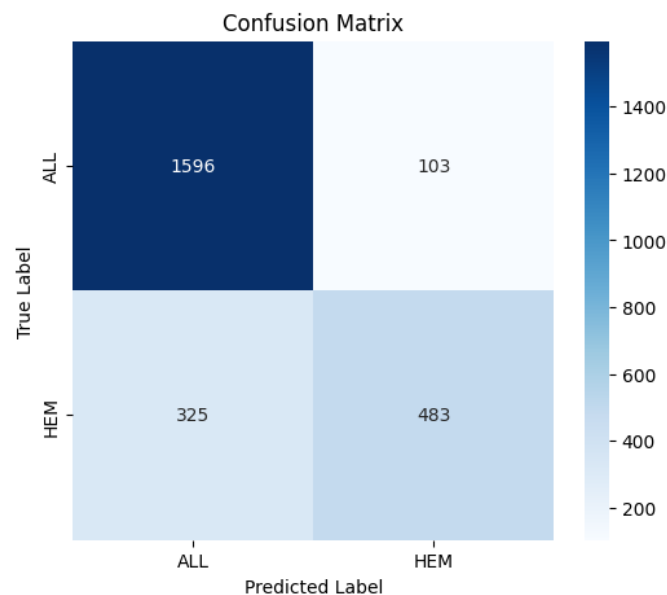
Pada pengujian dengan *batch size* 32, model dilatih selama 50 epoch menggunakan 10.021 citra latih dan diuji pada 2.507 citra uji. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.9.



Gambar 4. 9 Grafik Pelatihan Model Skenario 2 Batch Size 32

Berdasarkan Gambar 4.9, kurva training accuracy dan validation accuracy menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, training accuracy dimulai dari 69,05% dan validation accuracy dari 77,74%, kemudian keduanya meningkat secara bertahap hingga mencapai training accuracy sebesar 82,58% dan validation accuracy sebesar 82,93% pada epoch ke-50. Kedua kurva relatif berdekatan sepanjang proses pelatihan, mengindikasikan bahwa model tidak mengalami overfitting secara

signifikan. Kurva training loss dan validation loss juga menunjukkan tren penurunan yang konsisten, di mana training loss turun dari 0,6024 pada epoch pertama menjadi 0,4159 pada epoch ke-50, sedangkan validation loss turun dari 0,5259 menjadi 0,4201. Kedua kurva menunjukkan pola konvergensi yang stabil pada epoch-epoch akhir, mengindikasikan bahwa model telah berhasil mempelajari pola fitur citra sel darah dengan baik. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada confusion matrix yang ditunjukkan pada Gambar 4.10.



Gambar 4. 10 Confusion Matrix Skenario 2 Batch Size 32

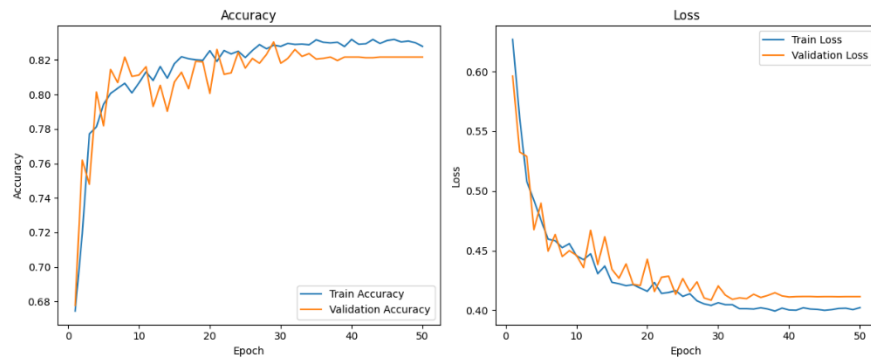
Berdasarkan Gambar 4.10, dari 1.699 citra ALL, sebanyak 1.596 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 103 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 808 citra HEM, sebanyak 483 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 325 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 32 menghasilkan akurasi sebesar 82,93% dengan *precision* kelas ALL sebesar

0,83, *recall* sebesar 0,94, dan *F1-score* sebesar 0,88. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,82, *recall* sebesar 0,60, dan *F1-score* sebesar 0,69.

Dibandingkan dengan *batch size* 16 pada Skenario 2, *batch size* 32 menghasilkan akurasi yang lebih rendah (82,93% vs 84,36%) dengan *recall* HEM yang juga lebih rendah (0,60 vs 0,63). Pola ini konsisten dengan hasil Skenario 1, di mana *batch size* yang lebih besar cenderung menghasilkan pembaruan bobot yang lebih jarang sehingga model membutuhkan lebih banyak iterasi untuk mencapai konvergensi optimal.

c. Batch Size 64

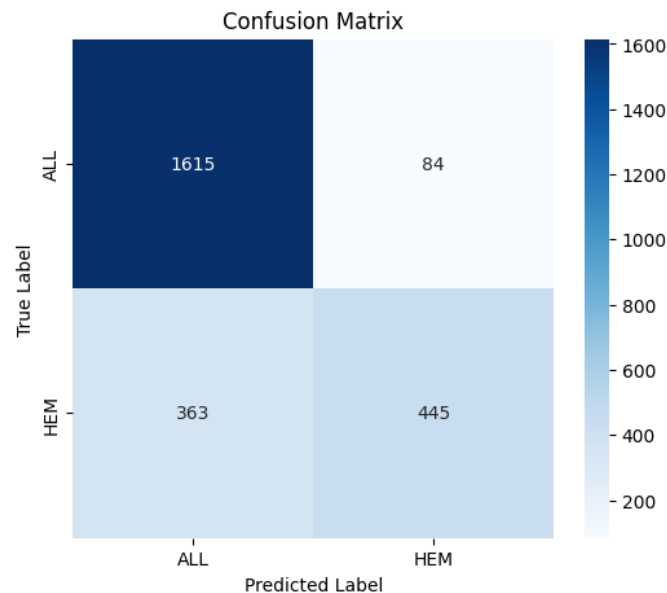
Pada pengujian dengan *batch size* 64, model dilatih selama 50 epoch menggunakan 10.021 citra latih dan diuji pada 2.507 citra uji. Dibandingkan dengan *batch size* 16 dan 32, *batch size* 64 menggunakan sampel yang lebih besar per iterasi sehingga estimasi gradien yang dihasilkan lebih stabil dan akurat. Dengan *batch size* 64 dan jumlah data latih sebesar 10.021 citra, model melakukan sebanyak 157 iterasi pembaruan bobot per epoch, lebih sedikit dibandingkan *batch size* 16 (627 iterasi) dan *batch size* 32 (314 iterasi). Meskipun jumlah iterasinya paling sedikit di antara ketiga variasi *batch size* pada Skenario 2, setiap iterasi menggunakan sampel yang lebih representatif dari keseluruhan distribusi data sehingga pembaruan bobot yang dilakukan lebih mencerminkan pola data secara keseluruhan. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.11.



Gambar 4. 11 Grafik Pelatihan Model Skenario 2 Batch Size 64

Berdasarkan Gambar 4.11, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 67,44% dan *validation accuracy* dari 67,77%, kemudian keduanya meningkat secara bertahap hingga mencapai *training accuracy* sebesar 82,79% dan *validation accuracy* sebesar 82,17% pada epoch ke-50. Dibandingkan *batch size* 16 dan 32, kurva *batch size* 64 menunjukkan peningkatan yang lebih lambat pada epoch-epoch awal, hal ini wajar karena *batch size* yang lebih besar menyebabkan pembaruan bobot yang lebih jarang per epoch sehingga model membutuhkan lebih banyak iterasi untuk konvergen. Kedua kurva relatif berdekatan sepanjang proses pelatihan, mengindikasikan bahwa model tidak mengalami *overfitting* secara signifikan. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,6269 pada epoch pertama menjadi 0,4023 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5963 menjadi 0,4115. Kedua kurva menunjukkan pola konvergensi yang stabil pada epoch-epoch akhir, mengindikasikan bahwa model telah berhasil mempelajari pola

fitur citra sel darah dengan baik. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.12.



Gambar 4. 12 Confusion Matrix Skenario 2 Batch Size 64

Berdasarkan Gambar 4.12, dari 1.699 citra ALL, sebanyak 1.615 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 84 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 808 citra HEM, sebanyak 445 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 363 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 64 menghasilkan akurasi sebesar 82,17% dengan *precision* kelas ALL sebesar 0,82, *recall* sebesar 0,95, dan *F1-score* sebesar 0,88. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,84, *recall* sebesar 0,55, dan *F1-score* sebesar 0,67.

Meskipun *batch size* 64 menghasilkan *recall* ALL yang sama tingginya dengan *batch size* 16 (0,95), nilai *recall* HEM yang paling rendah di antara ketiga variasi (0,55) menunjukkan bahwa model dengan *batch size* lebih besar

cenderung lebih kesulitan dalam mengklasifikasikan kelas HEM secara optimal. Hal ini konsisten dengan pola yang ditemukan pada Skenario 1, di mana ketidakseimbangan jumlah data antara kelas ALL dan HEM semakin terasa dampaknya pada *batch size* yang lebih besar.

d. Perbandingan Hasil Skenario 2

Setelah dilakukan serangkaian pengujian dengan tiga variasi *batch size*, yaitu *batch size* 16, 32, dan 64 pada Skenario 2 dengan rasio pembagian data 80:20, seluruh hasil evaluasi yang meliputi akurasi, *precision*, *recall*, dan *F1-score* dari masing-masing variasi *batch size* dirangkum secara keseluruhan pada Tabel 4.5 untuk memudahkan proses perbandingan dan analisis antar variasi.

Tabel 4. 5 Perbandingan Hasil Evaluasi Skenario 2 (80:20)

Batch Size	Accuracy	Precision	Recall	F1-Score
16	84,36%	84,5%	79%	80,5%
32	82,93%	82,5%	77%	78,5%
64	82,17%	83%	75%	77,5%

Berdasarkan Tabel 4.5, *batch size* 16 menghasilkan performa terbaik pada Skenario 2 dengan akurasi 84,36%, *precision* 84,5%, *recall* 79%, dan *F1-score* 80,5%. Pola ini konsisten dengan hasil Skenario 1, di mana *batch size* 16 juga menghasilkan performa terbaik. Meskipun *batch size* 64 menghasilkan *precision* yang cukup kompetitif (83%), nilai *recall* HEM yang paling rendah (0,55) menyebabkan *F1-score* keseluruhannya menjadi yang terendah di antara ketiga variasi (77,5%).

Secara keseluruhan, hasil Skenario 2 menunjukkan peningkatan performa yang konsisten dibandingkan Skenario 1 pada seluruh variasi *batch size*. Penambahan jumlah data latih dari 70% menjadi 80% terbukti memberikan

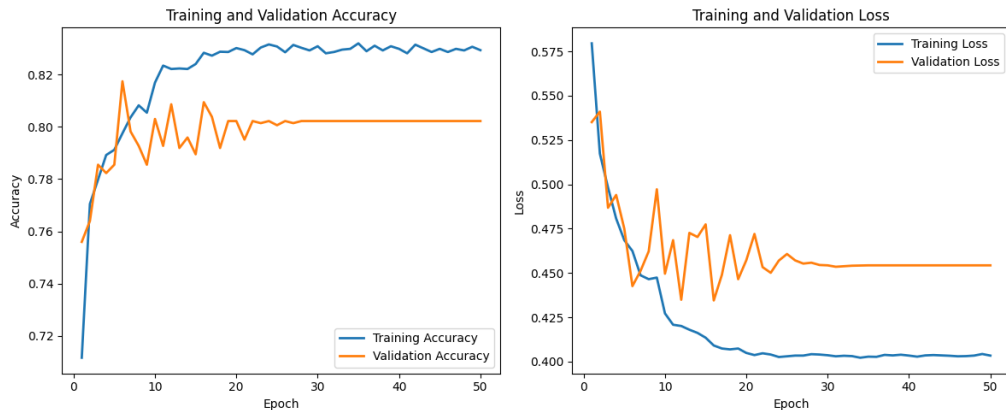
dampak positif terhadap kemampuan generalisasi model, yang terlihat dari peningkatan akurasi terbaik dari 83,16% pada Skenario 1 menjadi 84,36% pada Skenario 2. Hal ini mengindikasikan bahwa semakin banyak data latih yang digunakan, model semakin mampu mempelajari pola fitur yang lebih beragam sehingga meningkatkan kemampuannya dalam membedakan sel leukemia (ALL) dan sel darah normal (HEM). Peningkatan ini diharapkan akan berlanjut pada Skenario 3 dengan rasio pembagian data 90:10.

4.2.3 Hasil Uji Coba Skenario 3

Skenario 3 menggunakan pembagian data dengan rasio 90:10, di mana 90% data digunakan sebagai data latih dan 10% sebagai data uji. Dari total 12.528 citra, sebanyak 11.275 citra digunakan untuk pelatihan dan 1.254 citra digunakan untuk pengujian. Pada skenario ini, model dilatih selama 50 epoch dengan tiga variasi *batch size*, yaitu 16, 32, dan 64, untuk mengetahui pengaruh *batch size* terhadap performa model dengan jumlah data latih yang paling besar dibandingkan skenario sebelumnya.

a. Batch Size 16

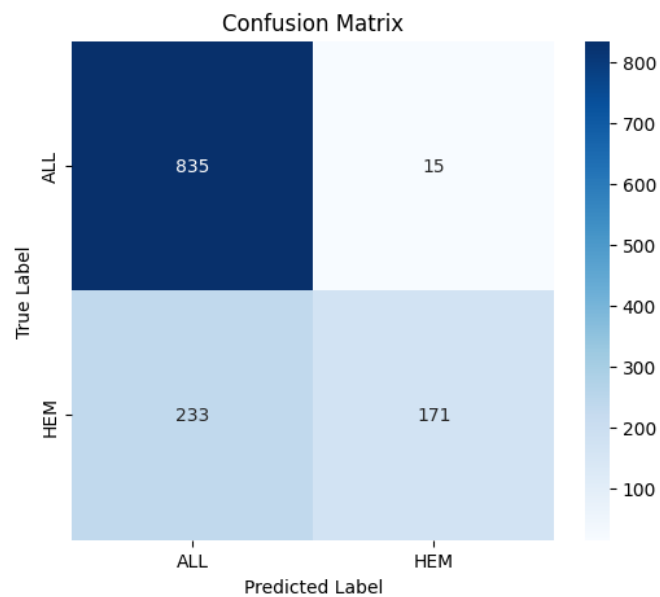
Pada pengujian dengan *batch size* 16, model dilatih selama 50 epoch menggunakan 11.275 citra latih dan diuji pada 1.254 citra uji. Dengan *batch size* 16 dan jumlah data latih sebesar 11.275 citra, model melakukan sebanyak 705 iterasi pembaruan bobot per epoch, yang merupakan jumlah iterasi terbanyak di antara seluruh variasi *batch size* pada Skenario 3. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.13.



Gambar 4. 13 Grafik Pelatihan Model Skenario 3 Batch Size 16

Berdasarkan Gambar 4.13, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 71,17% dan *validation accuracy* dari 75,60%, kemudian keduanya meningkat secara bertahap hingga mencapai *training accuracy* sebesar 82,93% dan *validation accuracy* sebesar 80,22% pada epoch ke-50. Terdapat kesenjangan yang cukup terlihat antara *training accuracy* dan *validation accuracy* pada epoch-epoch akhir, namun kedua kurva sudah mulai mendatar dan stagnan, mengindikasikan bahwa model telah mencapai titik konvergensi optimalnya. Stagnansi ini terjadi karena kombinasi jumlah data latih yang sangat besar dan *batch size* yang kecil menghasilkan jumlah iterasi per epoch yang sangat tinggi, sehingga model lebih cepat mencapai batas kapasitas belajarnya sesuai dengan arsitektur SqueezeNet yang digunakan. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,5795 pada epoch pertama menjadi 0,4034 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5351 menjadi 0,4543. Meskipun terdapat selisih

antara *training loss* dan *validation loss* pada epoch akhir, kedua kurva sudah menunjukkan pola yang stabil dan tidak berfluktuasi, mengonfirmasi bahwa model telah berhasil konvergen. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.14.



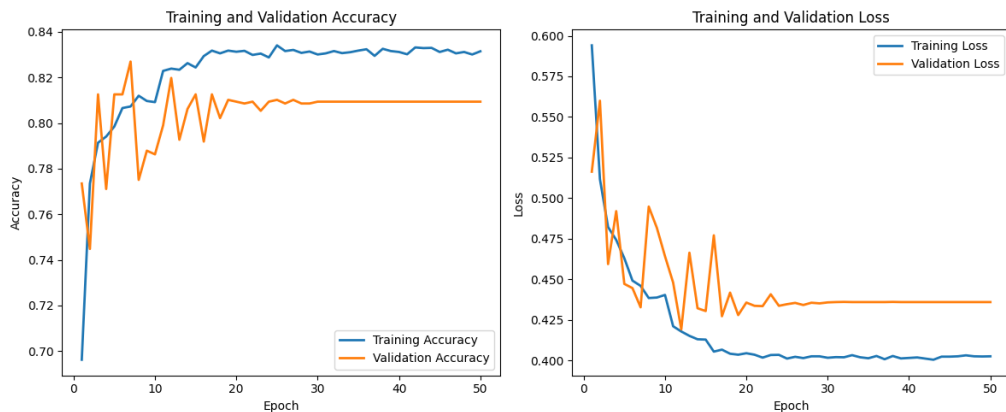
Gambar 4. 14 Confusion Matrix Skenario 3 Batch Size 16

Berdasarkan Gambar 4.14, dari 850 citra ALL, sebanyak 835 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 15 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 404 citra HEM, sebanyak 171 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 233 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 16 menghasilkan akurasi sebesar 80,22% dengan *precision* kelas ALL sebesar 0,78, *recall* sebesar 0,98, dan *F1-score* sebesar 0,87. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,92, *recall* sebesar 0,42, dan *F1-score* sebesar 0,58.

Hasil ini menunjukkan bahwa model dengan *batch size* 16 pada Skenario 3 sangat sensitif dalam mendeteksi sel leukemia, yang terlihat dari nilai *recall* ALL yang sangat tinggi (0,98). Namun di sisi lain, model ini menghasilkan nilai *recall* HEM yang paling rendah di antara seluruh variasi *batch size* pada Skenario 3 (0,42), yang berarti dari 404 citra HEM, sebanyak 233 citra (57,7%) salah diklasifikasikan sebagai ALL. Tingginya bias terhadap kelas ALL ini disebabkan oleh ketidakseimbangan jumlah data antara kelas ALL dan HEM, di mana dengan jumlah iterasi per epoch yang sangat tinggi (705 iterasi), model semakin cepat dan kuat dalam mempelajari pola kelas mayoritas (ALL) sehingga kemampuannya dalam mengenali kelas minoritas (HEM) menjadi kurang optimal. Meskipun demikian, *precision* HEM yang tinggi (0,92) menunjukkan bahwa ketika model memprediksi HEM, prediksi tersebut hampir selalu benar, hanya saja jumlah prediksi HEM yang dihasilkan sangat sedikit.

b. Batch Size 32

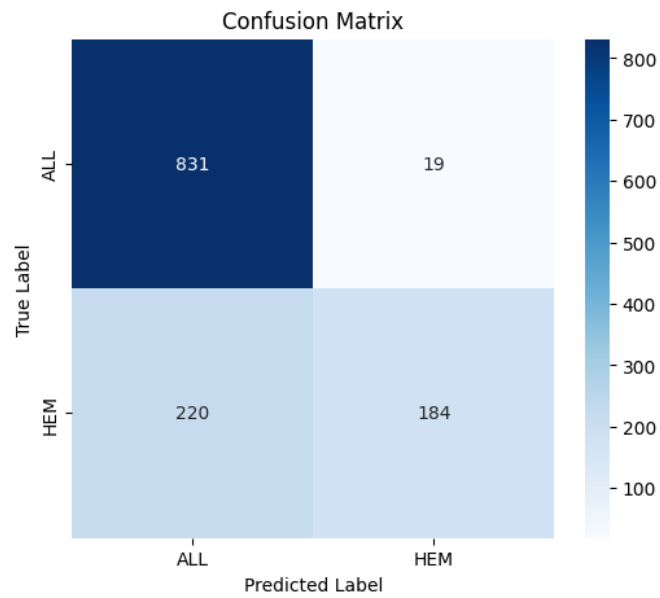
Pada pengujian dengan *batch size* 32, model dilatih selama 50 epoch menggunakan 11.275 citra latih dan diuji pada 1.254 citra uji. Dengan *batch size* 32 dan jumlah data latih sebesar 11.275 citra, model melakukan sebanyak 353 iterasi pembaruan bobot per epoch, lebih sedikit dibandingkan *batch size* 16 namun masih cukup tinggi untuk memastikan model belajar secara efektif dari seluruh data latih. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.15.



Gambar 4. 15 Grafik Pelatihan Model Skenario 3 Batch Size 32

Berdasarkan Gambar 4.15, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 69,62% dan *validation accuracy* dari 77,35%, kemudian keduanya meningkat secara bertahap hingga mencapai *training accuracy* sebesar 83,15% dan *validation accuracy* sebesar 80,94% pada epoch ke-50. Serupa dengan *batch size* 16, kedua kurva pada epoch-epoch akhir mulai mendatar dan stagnan, mengindikasikan bahwa model telah mencapai titik konvergensi optimalnya. Meskipun masih terdapat kesenjangan antara *training accuracy* dan *validation accuracy* pada epoch akhir, kesenjangan ini lebih kecil dibandingkan *batch size* 16, menunjukkan bahwa pengurangan jumlah iterasi per epoch memberikan dampak positif terhadap kemampuan generalisasi model. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,5941 pada epoch pertama menjadi 0,4026 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5163 menjadi 0,4360. Kedua kurva menunjukkan pola yang stabil pada epoch-epoch akhir, mengonfirmasi bahwa

model telah berhasil konvergen dengan baik. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.16.



Gambar 4. 16 Confusion Matrix Skenario 3 Batch Size 32

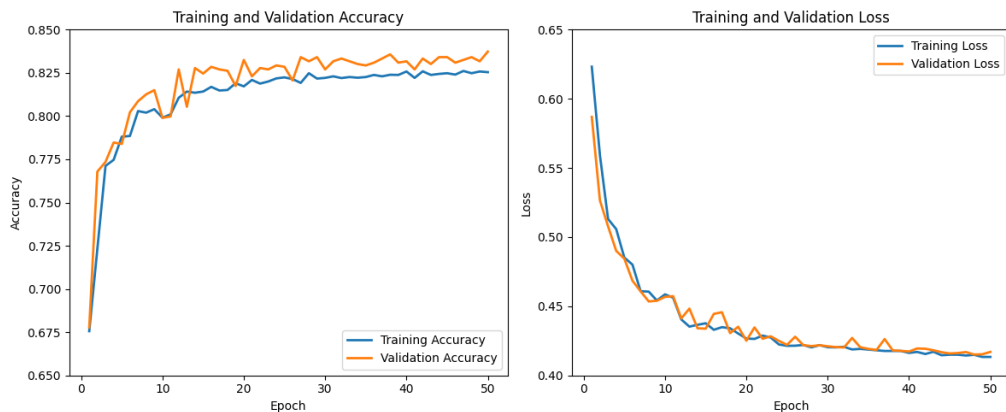
Berdasarkan Gambar 4.16, dari 850 citra ALL, sebanyak 831 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 19 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 404 citra HEM, sebanyak 184 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 220 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size* 32 menghasilkan akurasi sebesar 80,94% dengan *precision* kelas ALL sebesar 0,79, *recall* sebesar 0,98, dan *F1-score* sebesar 0,87. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,91, *recall* sebesar 0,46, dan *F1-score* sebesar 0,61.

Dibandingkan dengan *batch size* 16, *batch size* 32 menunjukkan peningkatan performa yang cukup berarti pada kelas HEM, di mana *recall* HEM meningkat

dari 0,42 menjadi 0,46 dan *F1-score* HEM meningkat dari 0,58 menjadi 0,61. Peningkatan ini mengindikasikan bahwa dengan jumlah iterasi per epoch yang lebih sedikit (353 iterasi vs 705 iterasi), model tidak terlalu agresif dalam mempelajari pola kelas ALL sehingga kemampuannya dalam mengenali kelas HEM sedikit lebih baik. Dari 404 citra HEM, sebanyak 220 citra (54,5%) masih salah diklasifikasikan sebagai ALL, lebih rendah dibandingkan *batch size* 16 (57,7%). Meskipun demikian, bias terhadap kelas ALL masih cukup signifikan, yang menunjukkan bahwa *batch size* 32 belum cukup besar untuk sepenuhnya mengatasi masalah ketidakseimbangan kelas pada Skenario 3.

c. Batch Size 64

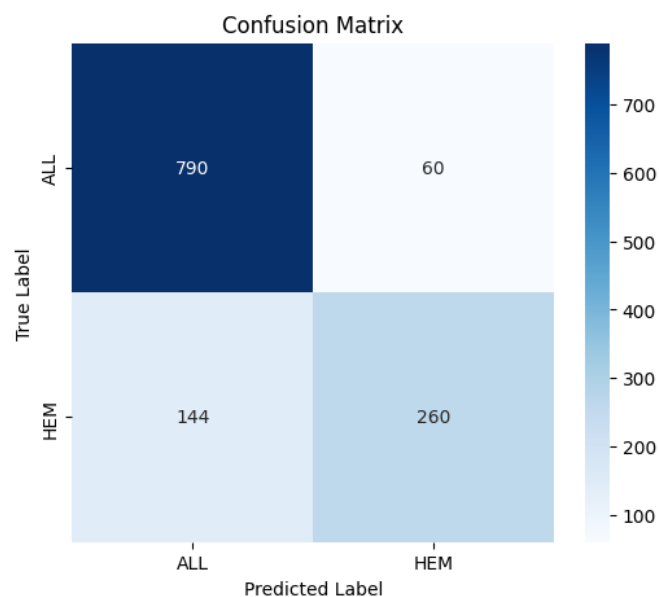
Pada pengujian dengan *batch size* 64, model dilatih selama 50 epoch menggunakan 11.275 citra latih dan diuji pada 1.254 citra uji. Dengan *batch size* 64 dan jumlah data latih sebesar 11.275 citra, model melakukan sebanyak 176 iterasi pembaruan bobot per epoch, paling sedikit di antara ketiga variasi *batch size* pada Skenario 3. Berkurangnya jumlah iterasi per epoch ini memberikan dampak yang signifikan terhadap cara model mempelajari pola fitur dari data latih, sebagaimana akan terlihat pada hasil evaluasi berikut. Proses pelatihan menghasilkan kurva *training accuracy*, *validation accuracy*, *training loss*, dan *validation loss* yang dapat dilihat pada Gambar 4.17.



Gambar 4. 17 Grafik Pelatihan Model Skenario 3 Batch Size 64

Berdasarkan Gambar 4.17, kurva *training accuracy* dan *validation accuracy* menunjukkan tren peningkatan yang konsisten dari epoch 1 hingga epoch 50. Pada epoch pertama, *training accuracy* dimulai dari 67,56% dan *validation accuracy* dari 67,78%, kemudian keduanya meningkat secara bertahap hingga mencapai *training accuracy* sebesar 82,54% dan *validation accuracy* sebesar 83,73% pada epoch ke-50. Berbeda dengan *batch size* 16 dan 32, terdapat fenomena menarik pada *batch size* 64 di mana kurva *validation accuracy* justru lebih tinggi dibandingkan *training accuracy* pada epoch-epoch akhir. Hal ini mengindikasikan bahwa model berhasil melakukan generalisasi yang sangat baik pada data uji tanpa mengalami *overfitting*, yang merupakan kondisi ideal dalam pelatihan model *deep learning*. Kedua kurva juga menunjukkan pola yang stabil dan mulai mendatar pada epoch-epoch akhir, mengonfirmasi bahwa model telah mencapai titik konvergensi yang optimal. Kurva *training loss* dan *validation loss* juga menunjukkan tren penurunan yang konsisten, di mana *training loss* turun dari 0,6232 pada epoch pertama menjadi 0,4134 pada epoch ke-50, sedangkan *validation loss* turun dari 0,5869 menjadi 0,4170.

Menariknya, *validation loss* pada epoch ke-50 (0,4170) lebih rendah dari *training loss* (0,4134), yang semakin mengonfirmasi kemampuan generalisasi model yang sangat baik pada *batch size* 64. Hal ini terjadi karena dengan jumlah iterasi per epoch yang lebih sedikit (176 iterasi), setiap pembaruan bobot menggunakan sampel yang lebih representatif dari keseluruhan distribusi data, sehingga model belajar pola yang lebih umum dan tidak terlalu terpaku pada karakteristik spesifik data latih. Distribusi hasil prediksi model secara lebih rinci dapat dilihat pada *confusion matrix* yang ditunjukkan pada Gambar 4.18.



Gambar 4. 18 Confusion Matrix Skenario 3 Batch Size 64

Berdasarkan Gambar 4.18, dari 850 citra ALL, sebanyak 790 citra berhasil diklasifikasikan dengan benar (*True Positive*) dan 60 citra salah diklasifikasikan sebagai HEM (*False Negative*). Dari 404 citra HEM, sebanyak 260 citra berhasil diklasifikasikan dengan benar (*True Negative*) dan 144 citra salah diklasifikasikan sebagai ALL (*False Positive*). Model dengan *batch size*

64 menghasilkan akurasi sebesar 83,73% dengan *precision* kelas ALL sebesar 0,85, *recall* sebesar 0,93, dan *F1-score* sebesar 0,89. Sementara itu, kelas HEM memperoleh *precision* sebesar 0,81, *recall* sebesar 0,64, dan *F1-score* sebesar 0,72.

Hasil ini menunjukkan peningkatan yang sangat signifikan dibandingkan *batch size* 16 dan 32 pada Skenario 3. *Recall* HEM meningkat drastis dari 0,42 (*batch size* 16) dan 0,46 (*batch size* 32) menjadi 0,64, sedangkan *F1-score* HEM meningkat dari 0,58 dan 0,61 menjadi 0,72. Dari 404 citra HEM, hanya 144 citra (35,6%) yang salah diklasifikasikan sebagai ALL, jauh lebih baik dibandingkan *batch size* 16 (57,7%) dan *batch size* 32 (54,5%). Peningkatan ini terjadi karena dengan jumlah iterasi per epoch yang lebih sedikit namun menggunakan sampel yang lebih besar dan lebih representatif per *batch*, model berhasil mempelajari pola fitur kedua kelas secara lebih seimbang. Fenomena ini membuktikan bahwa pada dataset dengan jumlah data latih yang sangat besar seperti Skenario 3, *batch size* yang lebih besar justru memberikan keunggulan dalam menghasilkan model yang lebih general dan tidak bias terhadap kelas mayoritas.

d. Perbandingan Hasil Skenario 3

Setelah dilakukan serangkaian pengujian dengan tiga variasi *batch size*, yaitu *batch size* 16, 32, dan 64 pada Skenario 3 dengan rasio pembagian data 90:10, seluruh hasil evaluasi yang meliputi akurasi, *precision*, *recall*, dan *F1-score* dari masing-masing variasi *batch size* dirangkum secara keseluruhan pada

Tabel 4.6 berikut untuk memudahkan proses perbandingan dan analisis antar variasi.

Tabel 4. 6 Perbandingan Hasil Evaluasi Skenario 3 (90:10)

Batch Size	Accuracy	Precision	Recall	F1-Score
16	80,22%	85%	70%	72,5%
32	80,94%	85%	72%	74%
64	83,73%	83%	78,5%	80,5%

Berdasarkan Tabel 4.6, terdapat pola yang berbeda dibandingkan Skenario 1 dan 2, di mana pada Skenario 3 justru *batch size* 64 yang menghasilkan performa terbaik dengan akurasi 83,73%, *precision* 83%, *recall* 78,5%, dan *F1-score* 80,5%. Sementara itu, *batch size* 16 dan 32 menghasilkan performa yang relatif lebih rendah dengan akurasi masing-masing 80,22% dan 80,94%, serta *recall* HEM yang sangat rendah yaitu 0,42 dan 0,46.

Perbedaan pola ini dapat dijelaskan melalui hubungan antara jumlah data latih dan jumlah iterasi per epoch pada masing-masing *batch size*. Pada Skenario 3 dengan 11.275 citra latih, *batch size* 16 menghasilkan 705 iterasi per epoch, *batch size* 32 menghasilkan 353 iterasi per epoch, sedangkan *batch size* 64 hanya menghasilkan 176 iterasi per epoch. Jumlah iterasi yang terlalu tinggi pada *batch size* 16 dan 32 menyebabkan model terlalu sering melakukan pembaruan bobot berdasarkan sampel yang relatif kecil, sehingga model menjadi terlalu agresif dalam mempelajari pola kelas mayoritas (ALL) dan kurang mampu menggeneralisasi pola kelas minoritas (HEM). Sebaliknya, *batch size* 64 dengan jumlah iterasi yang lebih sedikit namun menggunakan sampel yang lebih besar dan lebih representatif per *batch* berhasil

menghasilkan estimasi gradien yang lebih stabil, sehingga model dapat mempelajari pola kedua kelas secara lebih seimbang.

Secara keseluruhan, hasil Skenario 3 menunjukkan bahwa pengaruh *batch size* terhadap performa model tidak bersifat absolut, melainkan sangat bergantung pada jumlah data latih yang digunakan. Pada Skenario 3 dengan jumlah data latih terbesar (90%), *batch size* yang lebih besar terbukti lebih unggul dalam menghasilkan model yang lebih general dan tidak bias terhadap kelas mayoritas. Hal ini berbeda dengan Skenario 1 dan 2 di mana *batch size* kecil (16) lebih unggul karena pembaruan bobot yang lebih sering membantu model belajar lebih efektif pada jumlah data latih yang lebih sedikit.

4.3 Pembahasan

Setelah dilakukan serangkaian pengujian pada ketiga skenario dengan variasi rasio pembagian data dan *batch size*, bagian ini membahas analisis menyeluruh terhadap hasil yang diperoleh, meliputi pengaruh rasio pembagian data, pengaruh *batch size*, serta penentuan konfigurasi terbaik secara keseluruhan.

4.3.1 Pengaruh Rasio Pembagian Data

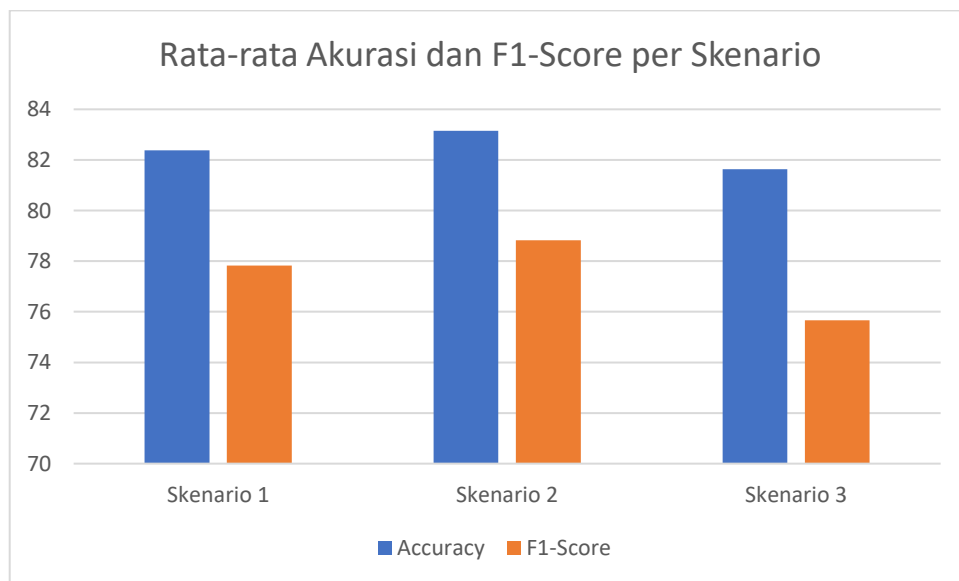
Salah satu variabel utama yang diuji dalam penelitian ini adalah rasio pembagian data latih dan data uji. Tiga variasi rasio yang digunakan adalah 70:30 pada Skenario 1, 80:20 pada Skenario 2, dan 90:10 pada Skenario 3. Pemilihan ketiga variasi rasio ini bertujuan untuk mengetahui sejauh mana jumlah data latih yang digunakan berpengaruh terhadap kemampuan model dalam mempelajari pola fitur citra sel darah dan melakukan generalisasi terhadap data yang belum pernah

dilihat sebelumnya. Untuk membandingkan pengaruh rasio pembagian data secara lebih objektif, perbandingan dilakukan berdasarkan nilai rata-rata akurasi dan rata-rata F1-Score dari ketiga variasi batch size pada masing-masing skenario, sehingga hasil perbandingan tidak dipengaruhi oleh faktor batch size tertentu. Perbandingan rata-rata akurasi dan F1-Score ketiga skenario dirangkum pada Tabel 4.7.

Tabel 4. 7 Rata-rata Akurasi dan F1-Score per Skenario

Skenario	Rasio Data	Rata-rata Akurasi	Rata-rata F1-Score
1	70:30	82,38%	77,83%
2	80:20	83,15%	78,83%
3	90:10	81,63%	75,67%

Berdasarkan Tabel 4.7, perbandingan rata-rata akurasi dan F1-Score ketiga skenario dapat divisualisasikan pada Gambar 4.19.



Gambar 4. 19 Grafik Perbandingan Rata-rata Akurasi per Skenario

Berdasarkan Tabel 4.7 dan Gambar 4.19, Skenario 1 dengan rasio 70:30 menghasilkan rata-rata akurasi sebesar 82,38% dan rata-rata F1-Score sebesar 77,83%. Pada skenario ini, model dilatih menggunakan 8.768 citra latih dan diuji pada 3.760 citra uji. Meskipun jumlah data latih yang digunakan relatif lebih sedikit

dibandingkan skenario lainnya, model sudah mampu mempelajari pola fitur dasar yang diperlukan untuk membedakan sel leukemia (ALL) dan sel darah normal (HEM) dengan cukup baik. Namun, keterbatasan jumlah data latih menyebabkan model belum sepenuhnya optimal dalam menggeneralisasi pola kelas HEM yang memiliki jumlah data lebih sedikit dibandingkan kelas ALL, yang tercermin dari nilai F1-Score yang masih belum optimal karena adanya ketidakseimbangan performa antara kedua kelas.

Pada Skenario 2 dengan rasio 80:20, rata-rata akurasi meningkat menjadi 83,15% dan rata-rata F1-Score meningkat menjadi 78,83%, keduanya menjadi yang tertinggi di antara ketiga skenario. Pada skenario ini, model dilatih menggunakan 10.021 citra latih dan diuji pada 2.507 citra uji. Penambahan jumlah data latih sebesar 1.253 citra dibandingkan Skenario 1 memberikan model lebih banyak variasi pola fitur untuk dipelajari, sehingga kemampuan generalisasi model terhadap data uji meningkat secara konsisten pada kedua metrik evaluasi. Peningkatan rata-rata akurasi dari 82,38% menjadi 83,15% serta peningkatan rata-rata F1-Score dari 77,83% menjadi 78,83% mengonfirmasi bahwa penambahan proporsi data latih dari 70% menjadi 80% memberikan dampak positif yang nyata terhadap performa model secara keseluruhan, baik dari sisi ketepatan klasifikasi maupun keseimbangan performa antar kelas.

Namun demikian, peningkatan rasio data latih dari 80% menjadi 90% pada Skenario 3 justru tidak menghasilkan peningkatan performa yang diharapkan. Rata-rata akurasi Skenario 3 turun menjadi 81,63% dan rata-rata F1-Score turun menjadi 75,67%, keduanya lebih rendah dibandingkan Skenario 1 maupun Skenario 2,

meskipun jumlah data latih yang digunakan paling besar yaitu 11.275 citra. Penurunan F1-Score yang lebih signifikan (dari 78,83% menjadi 75,67%) dibandingkan penurunan akurasi (dari 83,15% menjadi 81,63%) mengindikasikan bahwa ketidakseimbangan performa antar kelas semakin memburuk pada Skenario 3, khususnya pada kemampuan model dalam mengklasifikasikan kelas HEM. Fenomena ini terjadi karena beberapa faktor. Pertama, dengan rasio data uji yang semakin kecil yaitu hanya 10% atau 1.254 citra, kemampuan data uji dalam merepresentasikan keseluruhan distribusi data menjadi berkurang sehingga hasil evaluasi menjadi kurang stabil. Kedua, ketidakseimbangan antara kelas ALL dan HEM semakin terasa dampaknya ketika data uji semakin sedikit, karena jumlah citra HEM pada data uji Skenario 3 hanya 404 citra dibandingkan 808 citra pada Skenario 2 dan 1.212 citra pada Skenario 1, sehingga setiap kesalahan klasifikasi pada kelas HEM memberikan dampak yang lebih besar terhadap nilai metrik evaluasi secara keseluruhan. Ketiga, arsitektur SqueezeNet yang digunakan merupakan model yang ringan dengan kapasitas parameter terbatas (1.117.896 parameter), sehingga model kemungkinan sudah mencapai batas kapasitas belajarnya sebelum dapat memanfaatkan seluruh data latih yang tersedia secara optimal.

Berdasarkan analisis di atas, dapat disimpulkan bahwa penambahan jumlah data latih tidak selalu berbanding lurus dengan peningkatan performa model. Terdapat titik optimal dalam rasio pembagian data di mana keseimbangan antara jumlah data latih yang cukup untuk model belajar secara optimal dan jumlah data uji yang cukup representatif untuk mengevaluasi performa model secara akurat

menjadi sangat penting. Pada penelitian ini, rasio 80:20 terbukti menjadi rasio pembagian data yang paling optimal untuk model SqueezeNet yang digunakan, menghasilkan rata-rata akurasi tertinggi sebesar 83,15% dan rata-rata F1-Score tertinggi sebesar 78,83% di antara ketiga skenario yang diuji.

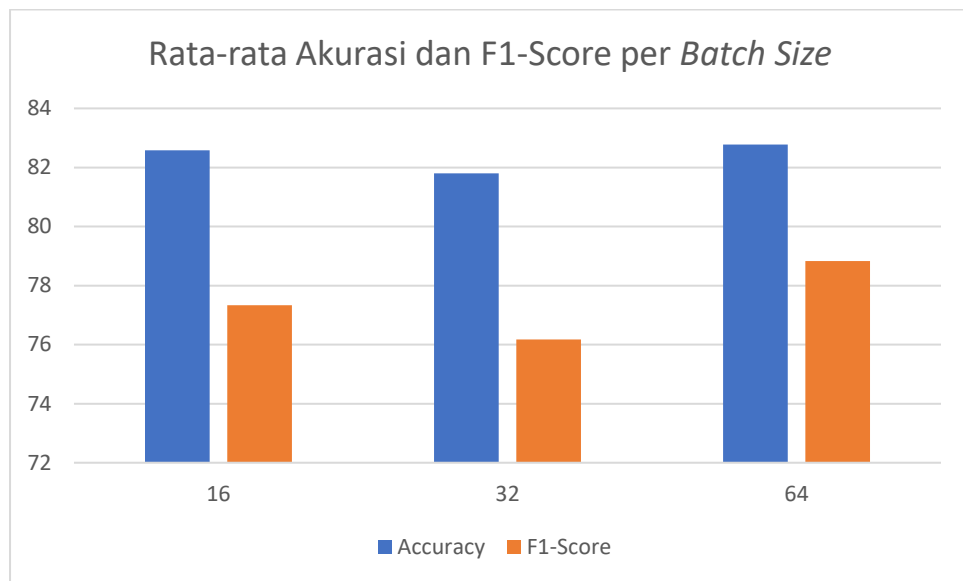
4.3.2 Pengaruh Batch Size

Selain rasio pembagian data, variabel lain yang diuji dalam penelitian ini adalah *batch size* yang digunakan selama proses pelatihan model. Tiga variasi *batch size* yang diuji adalah 16, 32, dan 64 pada masing-masing skenario. Untuk membandingkan pengaruh *batch size* secara lebih objektif, perbandingan dilakukan berdasarkan nilai rata-rata akurasi dan rata-rata F1-Score dari ketiga skenario pada masing-masing *batch size*, sehingga hasil perbandingan tidak dipengaruhi oleh faktor rasio pembagian data tertentu. Perbandingan rata-rata akurasi dan F1-Score ketiga variasi *batch size* dirangkum pada Tabel 4.8.

Tabel 4. 8 Rata-rata Akurasi dan F1-Score per Batch Size

Batch Size	Rata-rata Akurasi	Rata-rata F1-Score
16	82,58%	77,33%
32	81,80%	76,17%
64	82,78%	78,83%

Berdasarkan Tabel 4.8, terlihat bahwa ketiga variasi *batch size* menghasilkan performa yang cukup bervariasi baik dari sisi rata-rata akurasi maupun rata-rata F1-Score. Untuk memberikan gambaran yang lebih jelas dan memudahkan perbandingan secara visual, perbandingan rata-rata akurasi dan F1-Score ketiga variasi *batch size* dapat divisualisasikan pada Gambar 4.20.



Gambar 4. 20 Grafik Perbandingan Rata-rata Akurasi dan F1-Score per Batch Size

Berdasarkan Tabel 4.8 dan Gambar 4.20, terdapat pola yang menarik pada pengaruh *batch size* terhadap performa model secara keseluruhan. Secara rata-rata, *batch size* 64 menghasilkan performa tertinggi dengan rata-rata akurasi 82,78% dan rata-rata F1-Score 78,83%, diikuti *batch size* 16 dengan rata-rata akurasi 82,58% dan rata-rata F1-Score 77,33%, serta *batch size* 32 dengan rata-rata akurasi terendah 81,80% dan rata-rata F1-Score 76,17%. Namun demikian, untuk memahami pengaruh *batch size* secara lebih mendalam, perlu dilakukan analisis terhadap masing-masing variasi *batch size* pada setiap skenario secara terpisah sebagaimana diuraikan berikut.

a. Pengaruh Batch Size 16

Batch size 16 merupakan *batch size* terkecil yang diuji dalam penelitian ini.

Dengan *batch size* 16, model melakukan pembaruan bobot paling sering per epoch dibandingkan variasi *batch size* lainnya, yaitu 548 iterasi per epoch

pada Skenario 1, 627 iterasi per epoch pada Skenario 2, dan 705 iterasi per epoch pada Skenario 3.

Pada Skenario 1 dan Skenario 2, *batch size* 16 menghasilkan performa terbaik dibandingkan variasi *batch size* lainnya. Pada Skenario 1, *batch size* 16 menghasilkan akurasi 83,16% dengan F1-Score 79%, sedangkan pada Skenario 2 menghasilkan akurasi 84,36% dengan F1-Score 80,5%. Keunggulan ini disebabkan oleh frekuensi pembaruan bobot yang tinggi, yang memungkinkan model untuk lebih sering menyesuaikan bobotnya berdasarkan umpan balik dari data latih sehingga model dapat mempelajari pola fitur yang lebih beragam secara lebih efektif. Selain itu, *batch size* kecil juga memiliki efek regularisasi alami karena setiap pembaruan bobot didasarkan pada sampel yang lebih sedikit dan lebih beragam, sehingga model tidak terlalu bergantung pada pola spesifik dari satu kelompok sampel tertentu dan menghasilkan solusi yang lebih general.

Namun pada Skenario 3, *batch size* 16 justru menghasilkan performa yang paling rendah di antara ketiga variasi *batch size* dengan akurasi hanya 80,22% dan F1-Score 72,5%. Jumlah iterasi yang sangat tinggi (705 iterasi per epoch) dengan data latih yang sangat besar menyebabkan model melakukan pembaruan bobot yang terlalu sering dan terlalu agresif, sehingga model menjadi terlalu condong dalam mempelajari pola kelas mayoritas (ALL). Hal ini tercermin dari nilai *recall* HEM yang sangat rendah yaitu hanya 0,42, yang berarti lebih dari separuh citra HEM (57,7%)

salah diklasifikasikan sebagai ALL, sehingga berdampak signifikan terhadap penurunan F1-Score secara keseluruhan.

b. Pengaruh *Batch Size* 32

Batch size 32 merupakan variasi *batch size* menengah yang diuji dalam penelitian ini. Dengan *batch size* 32, model melakukan pembaruan bobot sebanyak 274 iterasi per epoch pada Skenario 1, 314 iterasi per epoch pada Skenario 2, dan 353 iterasi per epoch pada Skenario 3. Jumlah iterasi ini berada di antara *batch size* 16 dan 64, sehingga karakteristik pelatihan model dengan *batch size* 32 juga berada di antara keduanya.

Secara konsisten di seluruh skenario, *batch size* 32 menghasilkan performa yang berada di posisi tengah — tidak pernah menjadi yang terbaik maupun yang terburuk. Pada Skenario 1, *batch size* 32 menghasilkan akurasi 81,52% dengan F1-Score 76%. Pada Skenario 2, menghasilkan akurasi 82,93% dengan F1-Score 78,5%. Pada Skenario 3, menghasilkan akurasi 80,94% dengan F1-Score 74%. Pola ini menunjukkan bahwa *batch size* 32 memberikan keseimbangan yang cukup stabil antara frekuensi pembaruan bobot dan stabilitas estimasi gradien, meskipun tidak menghasilkan performa optimal pada kondisi manapun. Meskipun *batch size* 32 menghasilkan *recall* ALL yang tinggi pada Skenario 1 (0,96), hal ini disertai dengan *recall* HEM yang rendah (0,51), sehingga F1-Score keseluruhannya tetap lebih rendah dibandingkan *batch size* 16 pada skenario yang sama.

c. Pengaruh *Batch Size* 64

Batch size 64 merupakan *batch size* terbesar yang diuji dalam penelitian ini. Dengan *batch size* 64, model melakukan pembaruan bobot paling sedikit per epoch dibandingkan variasi *batch size* lainnya, yaitu 137 iterasi per epoch pada Skenario 1, 157 iterasi per epoch pada Skenario 2, dan 176 iterasi per epoch pada Skenario 3. Meskipun jumlah iterasinya paling sedikit, setiap iterasi pada *batch size* 64 menggunakan sampel yang lebih besar dan lebih representatif sehingga estimasi gradien yang dihasilkan lebih stabil dan akurat.

Pada Skenario 1 dan Skenario 2, *batch size* 64 menghasilkan performa yang berada di antara *batch size* 16 dan 32. Pada Skenario 1, *batch size* 64 menghasilkan akurasi 82,45% dengan F1-Score 78,5%, sedangkan pada Skenario 2 menghasilkan akurasi 82,17% dengan F1-Score 77,5%. Meskipun bukan yang terbaik pada kedua skenario ini, *batch size* 64 menunjukkan keseimbangan yang cukup baik antara performa kelas ALL dan HEM, dengan *recall* HEM yang lebih baik dibandingkan *batch size* 32 pada Skenario 1 (0,60 vs 0,51).

Pada Skenario 3, *batch size* 64 justru menghasilkan performa terbaik dengan akurasi 83,73% dan F1-Score 80,5%, melampaui *batch size* 16 dan 32 secara signifikan. Keunggulan ini disebabkan oleh kemampuan *batch size* 64 dalam menghasilkan estimasi gradien yang lebih stabil pada dataset yang sangat besar, sehingga model dapat mempelajari pola kedua kelas secara lebih seimbang. Hal ini terbukti dari nilai *recall* HEM yang meningkat

drastis menjadi 0,64 dibandingkan *batch size* 16 (0,42) dan *batch size* 32 (0,46), serta nilai *validation accuracy* yang justru lebih tinggi dari *training accuracy* pada epoch akhir, mengonfirmasi kemampuan generalisasi model yang sangat baik. Keunggulan *batch size* 64 pada Skenario 3 inilah yang menyebabkan rata-rata F1-Score *batch size* 64 secara keseluruhan (78,83%) menjadi yang tertinggi di antara ketiga variasi *batch size*.

Berdasarkan analisis di atas, dapat disimpulkan bahwa pengaruh *batch size* terhadap performa model bersifat dinamis dan sangat bergantung pada jumlah data latih yang digunakan. Pada jumlah data latih yang relatif lebih kecil seperti pada Skenario 1 dan Skenario 2, *batch size* kecil (16) lebih menguntungkan karena frekuensi pembaruan bobot yang tinggi membantu model belajar lebih efektif. Namun pada jumlah data latih yang sangat besar seperti pada Skenario 3, *batch size* yang lebih besar (64) justru lebih menguntungkan karena menghasilkan estimasi gradien yang lebih stabil dan model yang lebih general. Secara rata-rata keseluruhan, *batch size* 64 menghasilkan performa terbaik dengan rata-rata akurasi 82,78% dan rata-rata F1-Score 78,83%, menunjukkan bahwa pada dataset dengan karakteristik ketidakseimbangan kelas seperti yang digunakan dalam penelitian ini, *batch size* yang lebih besar cenderung menghasilkan model yang lebih seimbang dalam mengklasifikasikan kedua kelas secara keseluruhan.

4.3.3 Hasil Terbaik Keseluruhan

Setelah dilakukan analisis terhadap pengaruh rasio pembagian data dan *batch size* secara terpisah, bagian ini membahas penentuan konfigurasi terbaik secara keseluruhan berdasarkan seluruh hasil pengujian yang telah dilakukan.

Untuk memudahkan perbandingan, seluruh hasil pengujian dari kesembilan kombinasi skenario dan batch size dirangkum pada Tabel 4.9.

Tabel 4. 9 Perbandingan Hasil Seluruh Skenario dan Batch Size

Skenario	Rasio	Batch Size	Akurasi	Precision	Recall	F1-Score
1	70:30	16	83,16%	82,5%	77,5%	79%
1	70:30	32	81,52%	83%	73,5%	76%
1	70:30	64	82,45%	82%	76,5%	78,5%
2	80:20	16	84,36%	84,5%	79%	80,5%
2	80:20	32	82,93%	82,5%	77%	78,5%
2	80:20	64	82,17%	83%	75%	77,5%
3	90:10	16	80,22%	85%	70%	72,5%
3	90:10	32	80,94%	85%	72%	74%
3	90:10	64	83,73%	83%	78,5%	80,5%

Berdasarkan Tabel 4.9, hasil terbaik secara keseluruhan diperoleh pada Skenario 2 dengan *batch size* 16, yang menghasilkan akurasi 84,36%, *precision* 84,5%, *recall* 79%, dan *F1-score* 80,5%. Konfigurasi ini menggunakan rasio pembagian data 80:20 dengan 10.021 citra latih dan 2.507 citra uji, serta *batch size* 16 yang menghasilkan 627 iterasi pembaruan bobot per epoch.

Keunggulan konfigurasi Skenario 2 *batch size* 16 ini dapat ditinjau dari beberapa aspek. Pertama, dari sisi akurasi, konfigurasi ini menghasilkan akurasi tertinggi (84,36%) di antara seluruh kombinasi yang diuji, menunjukkan bahwa model mampu mengklasifikasikan citra sel darah dengan tingkat kebenaran yang paling tinggi. Kedua, dari sisi keseimbangan antar kelas, konfigurasi ini menghasilkan *recall* kelas ALL sebesar 0,95 dan *recall* kelas HEM sebesar 0,63, yang merupakan keseimbangan terbaik antara kemampuan mendeteksi sel leukemia dan kemampuan mengenali sel darah normal di antara seluruh konfigurasi yang diuji. Ketiga, dari sisi *F1-score*, konfigurasi ini menghasilkan *F1-score macro* sebesar 80,5% yang merupakan nilai tertinggi bersama dengan Skenario 3 *batch*

size 64, namun dengan akurasi yang lebih tinggi sehingga menjadikannya sebagai konfigurasi terbaik secara keseluruhan.

Perlu dicatat bahwa konfigurasi Skenario 3 dengan *batch size* 64 juga menghasilkan *F1-score* yang sama (80,5%), namun akurasinya lebih rendah (83,73% vs 84,36%). Meskipun *recall* HEM Skenario 3 *batch size* 64 sedikit lebih tinggi (0,64) dibandingkan Skenario 2 *batch size* 16 (0,63), keandalan hasil evaluasi Skenario 3 perlu dipertimbangkan dengan hati-hati mengingat jumlah data uji yang lebih sedikit (1.254 citra) dibandingkan Skenario 2 (2.507 citra), sehingga hasil evaluasi Skenario 2 dinilai lebih representatif dan dapat diandalkan sebagai tolak ukur performa model yang sebenarnya.

Dari perspektif klinis, hasil terbaik yang diperoleh pada penelitian ini memiliki implikasi yang penting. Nilai *recall* ALL sebesar 0,95 pada konfigurasi terbaik menunjukkan bahwa model mampu mendeteksi 95% kasus leukemia dengan benar, yang berarti hanya 5% kasus leukemia yang terlewat (*False Negative*). Dalam konteks medis, kemampuan mendeteksi kasus leukemia dengan *recall* yang tinggi sangat penting karena kegagalan dalam mendeteksi sel leukemia dapat berdampak serius terhadap penanganan pasien. Di sisi lain, nilai *recall* HEM sebesar 0,63 menunjukkan bahwa masih terdapat ruang untuk peningkatan dalam mengenali sel darah normal, yang dapat menjadi fokus pengembangan penelitian selanjutnya.

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa model SqueezeNet yang dimodifikasi dengan 10 *Fire Module* dan dikelompokkan ke dalam tiga level mampu mengklasifikasikan citra sel darah leukemia (ALL) dan sel

darah normal (HEM) dengan performa yang cukup baik. Konfigurasi terbaik yang diperoleh, yaitu Skenario 2 dengan *batch size* 16, menghasilkan akurasi 84,36% yang menunjukkan bahwa model ini berpotensi untuk digunakan sebagai alat bantu diagnosis leukemia berbasis citra sel darah, meskipun masih diperlukan pengembangan lebih lanjut untuk meningkatkan performa model terutama dalam mengklasifikasikan kelas HEM.

4.3.4 Perbandingan dengan Penelitian dan Arsitektur Terkait

Untuk mengevaluasi posisi model SqueezeNet yang dikembangkan dalam penelitian ini, dilakukan perbandingan dari dua aspek utama, yaitu perbandingan akurasi dengan penelitian yang menggunakan dataset C-NMC 2019, serta perbandingan efisiensi model dengan arsitektur *deep learning* ringan lainnya. Hasil perbandingan secara menyeluruh dirangkum pada Tabel 4.10.

Tabel 4. 10 Perbandingan dengan Model lain

No	Peneliti	Metode	Dataset	Akurasi	Parameter	Ukuran Model
1	Mondal et al. (2021)	Ensemble CNN (WEN-kappa)	C-NMC 2019	89,8%	>50 juta	>200 MB
2	Mondal et al. (2021)	MobileNet (individual)	C-NMC 2019	84,1%	4,2 juta	16 MB
3	(Tan & Le, 2019)	EfficientNet-B0	ImageNet	77,1%	5,3 juta	20 MB
4	Penelitian ini	SqueezeNet	C-NMC 2019	84,36%	1,1 juta	4,26 MB

Berdasarkan tabel di atas, perbandingan dapat ditinjau dari dua perspektif. Pertama, dari sisi akurasi pada dataset C-NMC 2019, SqueezeNet yang dikembangkan dalam penelitian ini mencapai akurasi 84,36%, melampaui MobileNet individual milik Mondal et al (2021) yang mencapai 84,1% pada dataset yang sama. Meskipun pendekatan *ensemble* CNN milik Mondal et al (2021) mencapai akurasi lebih tinggi yakni 89,8%, hal tersebut diperoleh dengan

menggabungkan tujuh arsitektur CNN sekaligus yang menghasilkan ukuran model lebih dari 200 MB dan kompleksitas komputasi yang sangat tinggi, sehingga tidak praktis untuk diimplementasikan pada perangkat medis dengan keterbatasan sumber daya.

Kedua, dari sisi efisiensi model, SqueezeNet menunjukkan keunggulan yang signifikan dibandingkan seluruh arsitektur pembanding. Dengan hanya 1,1 juta parameter dan ukuran model 4,26 MB, SqueezeNet jauh lebih ringan dibandingkan EfficientNet-B0 yang memiliki 5,3 juta parameter dengan ukuran ~20 MB (Tan & Le, 2019), maupun MobileNet individual yang memiliki 4,2 juta parameter dengan ukuran ~16 MB (Mondal et al., 2021).

Dengan demikian, hasil penelitian ini membuktikan bahwa SqueezeNet mampu menghasilkan akurasi yang kompetitif bahkan melampaui MobileNet pada dataset leukemia yang sama, sekaligus menjadi arsitektur yang paling ringan dan efisien di antara seluruh model pembanding. Keunggulan ganda ini menjadikan SqueezeNet pilihan yang paling tepat untuk diimplementasikan pada perangkat medis dengan keterbatasan sumber daya komputasi, seperti sistem *point-of-care* dan perangkat portabel di fasilitas kesehatan daerah terpencil.

Dalam perspektif Islam, ketelitian dan kehati-hatian dalam memproses informasi merupakan prinsip yang sangat ditekankan, termasuk dalam konteks pengambilan keputusan di bidang kesehatan. Hal ini sejalan dengan firman Allah SWT dalam QS. Al-Hujurat [49]:6:

يَا أَيُّهَا الَّذِينَ آمَنُوا إِنْ جَاءَكُمْ فَاسِقٌ بِنَبَأٍ فَتَبَيَّنُوا أَنْ تُصِيبُوا قَوْمًا بِجَهَالَةٍ فَتُصْحَبُوا عَلَىٰ مَا فَعَلْتُمْ لَنِمِينَ ﴿٦﴾

"Wahai orang-orang yang beriman! Jika seseorang yang fasik datang kepadamu membawa suatu berita, maka telitilah kebenarannya, agar kamu tidak

mencelakakan suatu kaum karena kebodohan (kecerobohan), yang akhirnya kamu menyesali perbuatanmu itu." (QS. Al-Hujurat [49]:6).

dalam Tafsir Al-Misbah (Vol. 13) menafsirkan ayat ini sebagai perintah untuk melakukan *tabayyun*, yaitu sikap selektif dan proporsional dalam memilah informasi yang dapat dipercaya dan penting sebelum dijadikan dasar pengambilan keputusan. Dalam konteks penelitian ini, prinsip *tabayyun* tersebut tercermin pada proses evaluasi model klasifikasi yang dilakukan secara sistematis melalui pengukuran akurasi, *precision*, *recall*, dan *F1-score*. Hasil terbaik yang diperoleh pada Skenario 2 dengan *batch size* 16 menunjukkan akurasi sebesar 84,36%, yang mencerminkan bahwa model telah mampu memproses informasi visual citra sel darah secara teliti dan dapat dipertanggungjawabkan. Dalam bidang medis, ketelitian prediksi model sangat krusial karena kesalahan klasifikasi dapat berdampak serius terhadap keselamatan pasien. Oleh karena itu, penelitian ini berupaya menghadirkan sistem klasifikasi yang akurat sebagai bentuk implementasi prinsip *tabayyun* dalam teknologi, yaitu menghasilkan informasi yang benar, terukur, dan dapat dijadikan landasan keputusan diagnosis yang tepat.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan, model SqueezeNet dengan 10 *Fire Module* terbukti mampu mengklasifikasikan citra sel darah *Acute Lymphoblastic Leukemia* (ALL) dan sel darah normal (HEM) menggunakan dataset C-NMC 2019 sebagai pendukung diagnosis berbasis citra. Hasil terbaik diperoleh pada skenario pembagian data 80:20 dengan *batch size* 16, yang menghasilkan akurasi sebesar 84,36%, *precision* sebesar 84,5%, *recall* sebesar 79%, dan *F1-score* sebesar 80,5%. Nilai *recall* kelas ALL sebesar 0,95 menunjukkan bahwa model mampu mendeteksi 95% kasus leukemia dengan benar, yang merupakan capaian penting dari perspektif klinis. Dengan demikian, model SqueezeNet terbukti memiliki performa yang cukup baik dalam mengklasifikasikan citra mikroskopis sel darah leukemia sehingga berpotensi diimplementasikan sebagai teknologi pendukung diagnosis *Acute Lymphoblastic Leukemia* berbasis citra sel darah.

5.2 Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat diberikan untuk pengembangan penelitian selanjutnya agar performa model klasifikasi dapat ditingkatkan dan menghasilkan hasil yang lebih optimal, yaitu sebagai berikut:

1. Penelitian selanjutnya disarankan untuk menerapkan tahap preprocessing yang lebih optimal, seperti segmentasi inti sel sebelum proses klasifikasi,

sehingga model dapat lebih fokus dalam mengekstraksi fitur yang relevan dari bagian sel yang paling informatif dan mengurangi pengaruh noise dari latar belakang citra

2. Mengingat model masih mengalami kesulitan dalam mengklasifikasikan kelas HEM dibandingkan kelas ALL, penelitian selanjutnya disarankan untuk menangani ketidakseimbangan kelas secara lebih eksplisit, misalnya dengan menerapkan teknik *weighted loss function* atau *oversampling* berbasis SMOTE, sehingga hasil klasifikasi dapat menjadi lebih akurat dan seimbang di kedua kelas.
3. Penelitian berikutnya dapat mencoba arsitektur CNN yang lebih modern seperti MobileNetV3 yang juga dirancang untuk efisiensi komputasi namun dengan kapasitas representasi fitur yang lebih baik, serta melakukan *tuning hyperparameter* secara lebih sistematis menggunakan pendekatan *Bayesian optimization* untuk memperoleh konfigurasi model yang lebih optimal.

DAFTAR PUSTAKA

- Ahmed, T. S., Philip, K., & Ahmed, A. E. (2020). *Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features*. 1–11. <https://doi.org/10.1038/s41598-020-59215-9>
- Akbar, M., Prasetyaningrum, P. T., Setyaningsih, P. W., & Ahsan, M. (2024). *Deteksi Leukemia Limfoblastik Akut menggunakan Convolutional Neural Network*. 7(1). <https://doi.org/10.32493/jtsi.v7i1.34168>
- Alifian, K., Wibowo, H., Azis, N., & Rianto, K. (2024). *Improved convolutional neural network model for leukemia classification using EfficientNetV2M and bayesian optimization*. 192–199.
- Baygin, M. (2022). *Vegetable and Fruit Image Classification with SqueezeNet based Deep Feature Generator*. 17(1), 121–134.
- Chang, J. H., Marcus, K. J., Poppe, M. M., Hua, C., & Esiashvili, N. (2021). *Acute lymphoblastic leukemia*. 68(February 2020), 1–7. <https://doi.org/10.1002/pbc.28371>
- Danendra, F. A., Anggraeny, F. T., & Hendra. (2023). *Klasifikasi Citra Penyakit Daun Cabai Rawit Dengan Menggunakan CNN Arsitektur AlexNet dan SqueezeNet*. 12(01), 50–61.
- Desiana, P., Ayu, W., & Pradipta, G. A. (2024). *SqueezeNet Feature Extraction dan Gradient Boosting untuk Klasifikasi Penyakit Monkeypox pada Citra Kulit*. 177–183.
- Elsayed, B., Elhadary, M., Salem, M. E., Khadadah, F., & Alshurafa, A. (2023). *Deep learning enhances acute lymphoblastic leukemia diagnosis and classification using bone marrow images*. December, 1–13. <https://doi.org/10.3389/fonc.2023.1330977>
- Fajrina, A. N., Pradana, Z. H., Purnama, S. I., & Romadhona, S. (2024). *Penerapan Arsitektur EfficientNet-B0 Pada Klasifikasi Leukimia Tipe Acute Lymphoblastik Leukimia Application of the EfficientNet-B0 Architecture in the Classification of Acute Lymphoblastic Leukemia*. 6(1).
- Gupta, R., Gehlot, S., & Gupta, A. (2022). *C-NMC : B-lineage acute lymphoblastic leukaemia : A blood cancer dataset*. *Medical Engineering and Physics*, 103(December 2021), 103793. <https://doi.org/10.1016/j.medengphy.2022.103793>
- Hao, M., Sun, Q., & Xuan, C. (2024). *applied sciences SqueezeNet : An Improved Lightweight Neural Network for Sheep Facial Recognition*.
- Harahap, H., Rahman, S., Zen, M., & Suriati. (2024). *Analisis Existing Convolutional Neural Network Untuk Klasifikasi Usia Pengunjung*. 4(1), 16–24.

- Huang, M., & Huang, Z. (2024). *lymphoblastic leukemia model for acute lymphoblastic leukemia image classification*. 21(October 2023), 1959–1978. <https://doi.org/10.3934/mbe.2024087>
- Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2017). *50 X FEWER PARAMETERS AND < 0.5 MB MODEL SIZE*. 1–13.
- Inaba, H., & Pui, C. (2021). *Advances in the Diagnosis and Treatment of Pediatric Acute Lymphoblastic Leukemia*.
- Istiqomah, H., & Ardianto, R. (2024). *Prediksi Kanker Darah Menggunakan Metode Convolutional Neural Network*. 4(1), 51–60.
- June, V. N., Putu, N., & Agustina, D. (2025). *JURNAL LOCUS: Penelitian & Pengabdian Klasifikasi Subtipe Leukemia Limfoblastik Akut (LLA) pada Citra Mikroskopis Sel Darah Menggunakan Arsitektur EfficientNet-B3 dengan Dataset Seimbang*. 4(6), 2978–2993.
- Minu, M. S., R, A. C., & S, S. R. S. (2022). *Traitement du Signal Optimal Squeeze Net with Deep Neural Network-Based Aerial Image Classification Model in Unmanned Aerial Vehicles*. 39(1), 275–281.
- Mondal, C., Hasan, K., Ahmad, M., Awal, A., Jawad, T., Dutta, A., Islam, R., & Ali, M. (2021). Informatics in Medicine Unlocked Ensemble of Convolutional Neural Networks to diagnose Acute Lymphoblastic Leukemia from microscopic images. *Informatics in Medicine Unlocked*, 27, 100794. <https://doi.org/10.1016/j.imu.2021.100794>
- Özcan, Ş. N., & Uyar, T. (2024). *Comprehensive data analysis of white blood cells with classification and segmentation by using deep learning approaches. November 2023*, 501–520. <https://doi.org/10.1002/cyto.a.24839>
- Pradana, A. G., Ignatius, D. R., Setiadi, M., Muslikh, A. R., Studi, P., Informatika, T., Komputer, F. I., Nuswantoro, U. D., Studi, P., Informasi, S., Informasi, F. T., & Malang, U. M. (2024). *Fine tuning model Convolutional Neural Network EfficientNet-B4 dengan augmentasi data untuk klasifikasi penyakit kakao Fine tuning Convolutional Neural Network model EfficientNet-B4 with data*. 2(1), 1–11. <https://doi.org/10.26905/jisad.v2i1.11899>
- Quiroga, F., Ronchetti, F., & Lanzarini, L. (2023). *Revisiting Data Augmentation for Rotational Invariance in Convolutional Neural Networks*.
- Rizki Firdaus Mulya, Ema Utami, D. A. (2026). *JURNAL RESTI Classification of Acute Lymphoblastic Leukemia based on White Blood*. 5(158), 947–952.
- Sampathila, N., Chadaga, K., Goswami, N., Chadaga, R. P., Pandya, M., Prabhu, S., Bairy, M. G., Katta, S. S., Bhat, D., & Upadya, S. P. (2022). *Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images*.
- Sekar, M. D., & Raj, M. (2023). *Role of Morphology in the Diagnosis of Acute*

Leukemias : Systematic Review. 464–473.

- Setiawan, W., Ghofur, A., Rachman, F. H., & Rulaningtyas, R. (2021). *Deep convolutional neural network alexnet and squeezenet for maize leaf diseases image classification*. 4, 0–7.
- Shahzad, M., Ali, F., Shirazi, S. H., Ahmad, A., Shah, B., & Kwak, D. (2024). *Blood cell image segmentation and classification : a systematic review*. 1–35. <https://doi.org/10.7717/peerj-cs.1813>
- Shihab, M. Q. (2002). *Tafsir Al Mishbah : pesan, kesan dan keserasian Al-Qur'an* (Cetakan I). Lentera Hati.
- Tan, M., & Le, Q. V. (2019). *EfficientNet : Rethinking Model Scaling for Convolutional Neural Networks*.
- Tavakoli, S., Ghaffari, A., Kouzehkhanan, Z. M., & Hosseini, R. (2021). New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images. *Scientific Reports*, 1–13. <https://doi.org/10.1038/s41598-021-98599-0>
- Ullah, M. Z., Zheng, Y., Song, J., Aslam, S., Xu, C., Kiazolu, G. D., & Wang, L. (2021). *applied sciences An Attention-Based Convolutional Neural Network for Acute Lymphoblastic Leukemia Classification*.
- Zhang, Y. (2023). *Optimized Convolutional Neural Networks for Enhanced Detection of Acute Lymphoblastic Leukemia from Grayscale Cell Images*. 571–577.