

BAB I

PENDAHULUAN

1.1 Latar Belakang

Leukemia limfoblastik akut (dalam bahasa Inggris dikenal dengan istilah *Acute Lymphoblastic Leukemia* atau ALL) merupakan proses mutasi sel progenitor limfoid yang menyebabkan sel darah putih memproduksi dirinya secara abnormal pada sumsum tulang, darah serta situs ekstraseluler lainnya, akibatnya sel darah putih normal mengalami disfungsi [1]. *American Cancer Society* memaparkan fakta bahwa sedikitnya 6660 kasus ALL terkonfirmasi di Amerika Serikat dengan angka kematian yang disebabkan yaitu 1560 kasus pada tahun 2022 yang menimpa orang dewasa dan anak-anak [2]. Sebesar 60% kasus ALL menimpa anak-anak yang berumur dibawah 15 tahun, sementara pada orang dewasa umumnya terjadi ketika mereka berusia lebih dari 50 tahun. ALL terbilang cukup jarang dialami orang dewasa, namun beberapa kondisi seperti bertambahnya usia, dapat meningkatkan resiko terjadinya ALL. Tercatat sekitar 20% kasus ALL menimpa orang dewasa [1]. Peningkatan kasus ALL setiap tahun cukup signifikan, dimana sebanyak 0.7% kasus yang telah terkonfirmasi menyerang anak-anak. Dapat disimpulkan bahwa ALL tergolong penyakit mematikan pada anak-anak yang harus segera mendapat penanganan. Rumah Sakit Kanker Dharmas memberikan informasi terkait peningkatan kasus kanker terutama pada anak di tahun 2006 hingga 2014. Puncaknya pada tahun 2014, kasus leukemia pada anak-anak mencapai 46 kasus sehingga menduduki posisi teratas dalam daftar penyebaran kanker pada anak [3].

Sejauh ini, belum dapat dipastikan apa penyebab terjadinya ALL pada seseorang. Namun beberapa faktor telah dikaitkan menjadi penyebab seseorang memiliki kemungkinan terkena ALL, diantaranya yaitu : menjalani pengobatan kemoterapi karena memiliki riwayat penyakit kanker sebelumnya, terkena paparan radiasi dengan intensitas sangat tinggi, kebiasaan merokok atau terpapar asap rokok, pestisida, dan kelainan genetik. Belum dapat dipastikan faktor-faktor tersebut dapat

menyebabkan ALL, namun penelitian dalam bidang tersebut masih tetap dilanjutkan [4].

ALL dapat di deteksi melalui metode pengujian mikroskopik pada *peripheral blood-smear* oleh hematologi. Pemeriksaan secara visual dilakukan dengan mengamati morfologi sel limfoblas untuk mengidentifikasi limfoblas pada sel darah yang dianalisis. Deteksi dini ALL perlu untuk dilakukan mengingat perkembangannya yang cepat dan dapat berakibat fatal jika tidak segera mendapatkan tindakan, hal ini juga berguna untuk menentukan pengobatan yang efektif. Sedangkan analisis morfologi dengan mikroskopis masih dilakukan secara manual oleh hematolog, sehingga memerlukan tenaga dan waktu yang cukup lama, hal ini juga bergantung akan keterampilan masing-masing hematolog [5].

Prosedur otomatis untuk mendukung pemeriksaan medis telah dilakukan sebelumnya, terutama untuk membantu skrining ALL menggunakan metode identifikasi dengan mengolah dan menganalisis bagian *peripheral blood-smear* (sediaan apus darah). Pada tahun 2019, penelitian untuk identifikasi ALL pada sel darah telah dilakukan. Berawal dengan inialisasi proses untuk mengidentifikasi sel darah menggunakan teknik *threshold*, yang kemudian dilanjutkan dengan ekstraksi ciri geometrik untuk klasifikasi menggunakan *Support Vector Machine (SVM) Linear* [6]. Hasil yang didapat menunjukkan bahwa metode yang diusulkan mampu mengklasifikasikan semua fitur yang diekstraksi dari sel darah putih; baik sel limfoblast atau ALL dan sel normal, dengan sensitivitas 98,4% dan tingkat kesalahan klasifikasi 6,1%. Pengembangan lebih lanjut perlu dilakukan, yaitu dengan memperhatikan proses dalam memisahkan sel darah yang tumpang tindih, yang sangat penting dalam menghitung semua sel individu darah putih pada citra sediaan apus darah.

Pada tahun 2021, penelitian untuk mengidentifikasi secara otomatis dan klasifikasi sel limfoblas dari gambar mikroskopis apusan darah tepi dilakukan dengan memperkenalkan kombinasi baru dari metodologi *image processing* dan mengusulkan *preprocessing* yang ekstensif dalam mencapai akurasi klasifikasi yang tinggi [7]. Secara khusus, kombinasi yang dipilih dari 16 fitur yang membawa informasi morfologi dan statistik menunjukkan kemampuan yang sangat baik untuk membedakan antara sel darah kanker dan non-kanker. Peneliti memilih sebagian

besar fitur berdasarkan kesamaannya dengan informasi visual, yang menjadi fokus pakar domain selama pemeriksaan manual. Fitur-fitur ini diekstraksi dari 241 *White Blood Cells* (WBC) yang tersegmentasi dari 31 gambar apusan darah tepi dari kumpulan data lokal. Metode yang diusulkan untuk melakukan klasifikasi adalah *Artificial Neural Network* (ANN) dan algoritma SVM. Model jaringan syaraf memberikan hasil yang lebih baik, mencapai sensitivitas 100% dan akurasi keseluruhan 97,52%. Probabilitas klasifikasi secara spesifik dari sel yang diidentifikasi dengan benar turut disajikan, juga telah dilakukan analisis terbalik untuk mengidentifikasi kegagalan klasifikasi penting. Pengamatan tersebut menunjukkan bahwa ketika akurasi yang dipublikasikan mencapai nilai tertinggi, metode klasifikasi mungkin tidak memberikan kejelasan atau keandalan yang cukup tinggi, dan oleh karena itu, pemeriksaan lebih lanjut diperlukan.

Pengembangan sistem untuk mendeteksi ALL secara otomatis kembali dilakukan pada tahun 2021. Tujuan utama dari penelitian tersebut yaitu untuk menyarankan fitur utama untuk mendeteksi ALL [8]. *Resize* dilakukan untuk mengubah semua gambar menjadi ukuran yang seragam. Inti dari sel limfosit dan sel limfoblas disegmentasi dengan menggunakan U-Net. Sekitar 1000 fitur mendalam diekstraksi dari lapisan yang terhubung penuh dari model *Convolutional Neural Network* (CNN) yang berbeda seperti AlexNet, GoogleNet, dan SqueezeNet. Penggabungan *deep feature* dilakukan untuk menggabungkan semua fitur. Fitur utama dipilih dengan menggunakan berbagai metode pemilihan fitur seperti *Recursive Feature Elimination* (RFE), *Mutual Information* (MI), dan *Minimum Recursive Maximal Relevance* (mRMR). Analisis statistik set fitur yang berbeda dilakukan dengan ANOVA. Hasil yang didapatkan adalah semua set *deep feature* yang menyatu adalah signifikan. Sebesar 50% dari *deep feature* yang bersimpangan menunjukkan signifikansi klinis yang konstruktif untuk menemukan sel-sel limfoblas secara tepat pada tahap awal.

Pada penelitian ini dilakukan klasifikasi penyakit ALL pada citra sel darah menggunakan model CNN arsitektur LeNet-5. Penggunaan arsitektur LeNet-5 yang dikombinasikan dengan metode CNN mampu mencapai tingkat akurasi sebesar 98%, meskipun hanya menggunakan fitur biner yang sederhana [9]. Pengujian juga dilakukan dengan menggunakan nilai *learning rate* dan *epoch* yang berbeda. Hal

itu bertujuan untuk mengetahui pengaruh *learning rate* dan epoch pada hasil akurasi yang dibuat oleh model LeNet-5 dalam mengklasifikasikan citra sel darah putih normal dan citra sel darah putih ALL.

1.2 Rumusan Masalah

Perumusan masalah yang dihasilkan dengan mengacu latar belakang adalah seperti berikut:

1.2.1 Permasalahan

Deteksi dini pada ALL sangat penting untuk dilakukan, mengingat perkembangannya yang cepat sehingga dibutuhkan langkah yang tepat dalam menentukan pengobatan yang efektif. Saat ini diagnosis pada ALL yang dilakukan oleh hematologi dengan menggunakan pemeriksaan visual dengan mikroskopis pada sel darah putih masih menggunakan metode manual. Pemeriksaan tersebut dilakukan untuk mengidentifikasi limfoblas pada sel darah yang dianalisis. Oleh karena pemeriksaan yang dilakukan masih secara manual, sehingga menimbulkan tenaga lebih dan waktu yang cukup lama, hal ini juga bergantung pada keterampilan dari para hematolog.

1.2.2 Solusi Permasalahan

Penelitian ini dilakukan untuk dapat mendeteksi serta mengidentifikasi ALL pada sel darah putih. Data yang digunakan adalah data citra sel darah putih. Proses *augmentasi* pada citra diterapkan pada data *training*. Data *training* hasil *augmentasi* dilatih untuk membentuk model hingga dapat mengklasifikasikan data *testing* menggunakan arsitektur LeNet-5. Pengklasifikasian tersebut akan memisahkan antara sel darah putih normal dan sel ALL.

1.2.3 Pertanyaan Penelitian

Berdasarkan latar belakang yang telah diuraikan di atas, dapat dirumuskan pertanyaan penelitian sebagai berikut:

1. Bagaimana pengaruh *learning rate* pada hasil akurasi yang dihasilkan model LeNet-5 dalam mengklasifikasikan citra sel darah putih dan sel ALL?
2. Bagaimana pengaruh *epoch* pada hasil akurasi yang dihasilkan model LeNet-5 dalam mengklasifikasikan citra sel darah putih dan sel ALL?
3. Berapa tingkat akurasi optimal yang dihasilkan dari model LeNet-5 sebagai metode klasifikasi pada citra sel darah putih dan sel ALL?

1.3 Tujuan dan Manfaat Penelitian

Penelitian ini mencakup beberapa tujuan dan manfaat yang dihasilkan, diantaranya yaitu:

1.3.1 Tujuan Penelitian

Penjabaran dari tujuan yang diperoleh pada penelitian ini adalah sebagai berikut:

1. Mengetahui pengaruh *learning rate* pada hasil akurasi yang dihasilkan model LeNet-5 dalam mengklasifikasikan citra sel darah putih dan sel ALL.
2. Mengetahui pengaruh *epoch* pada hasil akurasi yang dihasilkan model LeNet-5 dalam mengklasifikasikan citra sel darah putih dan sel ALL.
3. Mengetahui tingkat akurasi yang dihasilkan dari model LeNet-5 sebagai metode klasifikasi citra sel darah putih dan sel ALL.

1.3.2 Manfaat Penelitian

1. Memberikan pandangan pada studi baru di bidang penelitian terkait.
2. Dapat mendukung solusi baru untuk *image processing* dan komunitas *pattern matching*.
3. Dapat dijadikan sebagai bahan rujukan atau referensi untuk keperluan penelitian berikutnya.

1.4 Batasan Masalah

Penelitian ini menggunakan beberapa cakupan batasan masalah yang telah ditentukan, diantaranya yaitu :

1. *Dataset* yang digunakan untuk pelatihan dan pengujian menggunakan citra sel darah putih yang didapatkan dari Fabio Scotti yang berlokasi di Department of Information Technology - Universitas Degli Studi di Milano yaitu pada versi ALL_IDB2 [10]. Total data pada ALL_IDB2 adalah sebanyak 260 citra, dengan 2 kelas berbeda yaitu, kelas 0 untuk citra sel darah normal dan kelas 1 untuk sel ALL.
2. *Preprocessing* pada citra menggunakan teknik *augmentasi* data dengan penerapannya pada citra sebagai berikut; *Horizontal Flip*, *Vertical Flip*, *Random Rotation*, dan *Zoom Range*.
3. Nilai parameter yang digunakan pada *random rotation* adalah 30 derajat dan *zoom range* (0,7, 1,0) sehingga menghasilkan citra teraugmentasi dengan posisi *zoom in*.
4. Pengujian pada model menggunakan beberapa nilai *learning rate* yaitu : 0,01, 0,001, 0,0001, dan 0,00001 serta nilai *epoch* yang dijadikan uji coba yaitu 50, 100, dan 150.
5. Jumlah *batch size* pada penelitian ini adalah 32.

1.5 Sistematika Tugas Akhir

Pada penelitian ini, digunakan sistematika penulisan tugas akhir sebagai berikut:

BAB I PENDAHULUAN

Bab ini menguraikan penjelasan mengenai latar belakang dari masalah yang dibahas, perumusan masalah yang terjadi hingga solusi dari permasalahan, tujuan dan manfaat penerapan atas solusi yang diajukan, batasan-batasan, dan sistematika penulisan dalam tugas akhir ini.

BAB II KAJIAN PUSTAKA

Bab ini memuat kajian penelitian terdahulu yang melandasi penulisan penelitian ini, dan kajian pustaka yang memuat landasan teori seperti *Acute Lymphoblastic Leukemia*, *Augmentasi Citra*, *Convolutional Neural Network*, *Arsitektur LeNet-5*, *Keras*, *Evaluasi Model* dan *Penelitian Terkait*.

BAB III METODE PENELITIAN

Bab ini menjabarkan metode usulan pada penelitian yang mencakup jenis dan sumber data. Pada bab ini juga akan diuraikan teknik pengolahan data hingga langkah penelitian dan evaluasi hasil analisis klasifikasi.

BAB IV HASIL DAN PEMBAHASAN

Bab ini mejabarkan tentang hasil implementasi dari sistem dengan menggunakan metode arsitektur LeNet-5. Hasil dari implementasi terdiri dari yaitu: lingkungan uji coba, skenario uji coba, tahapan dalam pembuatan program, hasil uji coba, dan analisis hasil uji coba dari sistem.

BAB V PENUTUP

Bab ini membahas tentang kesimpulan yang didapat dari penelitian dengan menyesuaikan pada pertanyaan penelitian dan saran terkait pengembangan aplikasi selanjutnya.