

LEUKO-AI : SISTEM PENGECEMAN, PENGELASAN DAN PENGIRAAN SEL DARAH PUTIH MENGGUNAKAN TEKNIK PEMBELAJARAN MENDALAM

¹Nur Umairah Nabilah Binti Khairuddin,¹Afzan Binti Adam

**¹Fakulti Teknologi & Sains Maklumat, Universiti Kebangsaan Malaysia,
43600 UKM Bangi, Selangor Darul Ehsan, Malaysia**

Abstrak

Analisis sel darah putih merupakan aspek penting dalam mendiagnosis penyakit hematologi seperti leukemia, limfoma, jangkitan patogen dan gangguan imunisasi yang lain. Dalam era globalisasi ini, kaedah manual masih banyak digunakan dalam makmal perubatan. Namun, kaedah ini memerlukan masa yang lama, bergantung kepada kepakaran individu, dan terdedah kepada kesilapan manusia. Masalah ini boleh menjejaskan ketepatan diagnosis, terutamanya di makmal atau hospital yang kekurangan tenaga pakar. Walaupun terdapat usaha untuk mengautomatikkan proses ini, sistem sedia ada masih berdepan cabaran dalam memproses imej calitan darah yang kompleks, terutama apabila melibatkan kepelbagaian morfologi leukosit, kualiti imej yang tidak seragam, serta kewujudan sel yang bertindih. Faktor-faktor ini sering mengganggu ketepatan pengenalpastian dan pengelasan sel, sekali gus menyumbang kepada penurunan prestasi sistem apabila diaplikasikan dalam persekitaran klinikal sebenar. Di samping itu, komponen pengiraan automatik sel darah putih turut menghadapi masalah ketepatan, terutamanya apabila sel sukar dibezakan secara visual akibat pertindihan atau variasi intensiti pewarnaan pada imej. Sehubungan itu, kajian ini bertujuan untuk membangunkan LeukoAI, sebuah sistem pintar berasaskan pembelajaran mendalam yang mampu melaksanakan pengecaman dan pengelasan sel darah putih secara automatik daripada imej calitan darah, serta meningkatkan tahap ketepatan dan kelajuan dalam proses pengiraan sel darah putih secara menyeluruh dan efisien. Sistem ini dibangunkan menggunakan dua set data berbeza untuk melatih model pengecaman dan pengelasan secara khusus. Pengecaman sel darah putih dilaksanakan menggunakan model YOLOv8, yang dioptimumkan untuk pengecaman objek berskala kecil dan bertindih dalam imej perubatan. Bagi pengelasan lima jenis utama sel darah putih iaitu neutrofil, eosinofil, basofil, monosit dan limfosit, model DenseNet-121 digunakan kerana keupayaannya dalam mengekstrakan ciri mendalam dan generalisasi yang baik pada data perubatan. Proses pengiraan dijalankan secara automatik berdasarkan bilangan sel yang dikesan dalam setiap imej. Hasil eksperimen menunjukkan bahawa LeukoAI mencapai mAP50 melebihi 0.984, seperti yang diperoleh daripada graf Precision-Recall semasa penilaian model, menandakan ketepatan tinggi dalam pengecaman sel darah putih. Sistem turut

mencatat ketepatan pengecaman sebanyak 98%, membuktikan keupayaannya dalam mengurangkan kesilapan diagnosis dan mempercepatkan analisis. Bagi memastikan kebolehcapaian dan kemudahan penggunaan, LeukoAI akan diintegrasikan ke dalam platform web mesra pengguna. Antaramuka web ini membolehkan pengguna seperti juruteknik makmal dan pakar hematologi memuat naik imej calitan darah, serta memperoleh hasil pengecaman, pengelasan dan pengiraan secara automatik dalam masa nyata. Dengan pendekatan ini, LeukoAI berpotensi untuk diterapkan secara meluas dalam sistem diagnostik sebenar, terutamanya di makmal perubatan yang kekurangan kepakaran, dan seterusnya menyumbang kepada pembangunan teknologi diagnostik digital yang lebih pintar dan efisien.

Kata Kunci: Pembelajaran Mendalam, Sel Darah Putih, Pengecaman Automatik, YOLOv8, Diagnosis Hematologi.

Abstract

White blood cell analysis plays a vital role in diagnosing hematological diseases such as leukemia, lymphoma, pathogenic infections, and other immune disorders. In the current era of globalization, manual methods are still widely practiced in medical laboratories. However, these approaches are time-consuming, highly dependent on individual expertise, and prone to human error. This can compromise diagnostic accuracy, particularly in laboratories or hospitals that lack sufficient skilled personnel. Although efforts have been made to automate this process, existing systems still face challenges in processing complex blood smear images, especially due to the morphological diversity of white blood cells, inconsistent image quality, and overlapping cells. These factors often interfere with the accuracy of cell identification and classification, resulting in reduced system performance when applied in real clinical environments. Additionally, automatic white blood cell counting systems still suffer from accuracy issues, particularly when cells are visually difficult to distinguish due to overlapping or variations in staining intensity. In response to these challenges, this study aims to develop LeukoAI, an intelligent deep learning based system capable of automatically detecting and classifying white blood cells from blood smear images, while improving the accuracy and speed of the overall white blood cell counting process. The system is developed using two separate datasets to specifically train the detection and classification models. White blood cell detection is performed using the YOLOv8 model, which is optimized for detecting small and overlapping objects in medical images. For classifying the five major types of white blood cells, namely neutrophils, eosinophils, basophils, monocytes, and lymphocytes, the DenseNet121 model is used due to its deep feature extraction capabilities and strong generalization on medical data. The counting process is carried out automatically based on the number of detected cells in each image. Experimental results show that LeukoAI achieved a mean average precision at an intersection over union threshold of 0.5 (mAP50) exceeding 0.984, as obtained from the Precision Recall graph during model evaluation, indicating high accuracy in white blood cell detection. The system also recorded a detection accuracy of 98 percent,

demonstrating its effectiveness in reducing diagnostic errors and accelerating the analysis process. To ensure accessibility and user convenience, LeukoAI will be integrated into a user friendly web based platform. This web interface enables users such as laboratory technicians and hematologists to upload blood smear images and receive automatic results for detection, classification, and counting in real time. With this approach, LeukoAI holds strong potential for wide application in real diagnostic systems, especially in medical laboratories with limited expertise, and contributes to the advancement of smarter and more efficient digital diagnostic technologies.

Keywords : *Deep Learning, White Blood Cells, Automatic Recognition, YOLOv8, Hematology Diagnosis.*

1.0 PENGENALAN

Hematologi merupakan cabang ilmu perubatan yang memfokuskan kepada kajian darah, termasuk komponen-komponennya, fungsi fisiologi, dan pelbagai permasalahan yang berkaitan. Darah memainkan peranan penting dalam proses diagnostik klinikal kerana ia berfungsi sebagai medium utama untuk mengangkut oksigen dan nutrien ke seluruh tubuh. Di samping itu, darah juga menyumbang kepada kestabilan fisiologi melalui mekanisme seperti regenerasi sel, pembekuan darah, pengawalan suhu badan, dan perlindungan sistem imun (Thinam Tamang et al., 2022). Secara asasnya, darah terdiri daripada empat komponen utama iaitu sel darah merah, sel darah putih, platelet, dan plasma. Sel darah putih memainkan peranan penting dalam pertahanan imun badan dengan mengenal pasti dan memusnahkan patogen seperti bakteria, virus, sel abnormal, dan bahan asing lain (Ak et al., 2018). Leukosit boleh dikategorikan kepada dua kumpulan utama berdasarkan kehadiran granul dalam sitoplasma, iaitu granulosit yang terdiri daripada neutrofil, eosinofil dan basofil serta agranulosit iaitu limfosit dan monosit. Masing-masing dengan ciri morfologi dan fungsi imun yang berbeza.

Pengenalpastian jenis leukosit adalah aspek kritikal dalam menilai keadaan kesihatan pesakit, khususnya dalam mengenal pasti penyakit berkaitan sistem darah dan imun. Sebagai contoh, neutrofil bertindak sebagai barisan pertahanan pertama terhadap jangkitan bakteria, eosinofil terlibat dalam tindak balas alahan dan jangkitan parasit, basofil menghasilkan bahan kimia seperti histamin yang mencetuskan reaksi alahan, limfosit berperanan dalam imuniti adaptif melalui penghasilan antibodi, manakala monosit pula menjalankan fungsi fagositosis terhadap patogen dan sel mati. Dalam konteks amalan klinikal, Ujian Kiraan Darah Lengkap (CBC) merupakan prosedur diagnostik standard yang digunakan untuk menilai komponen darah, termasuk jumlah sel darah merah, sel darah putih dan platlet. Meskipun sel darah putih hanya mewakili kira-kira satu peratus daripada keseluruhan sel darah, peranannya dalam sistem imun menjadikannya komponen yang amat penting dalam proses diagnosis

(Cleveland Clinic, n.d.). Ujian ini biasanya disertakan dengan analisis pembezaan untuk mengenal pasti jenis dan bilangan khusus leukosit.

Namun begitu, kaedah manual menggunakan hemositometer masih digunakan secara meluas dalam makmal klinikal, seperti yang dilaporkan oleh Lee, Chen dan Lin (2022). Kaedah ini memerlukan kepakaran teknikal yang tinggi dan terdedah kepada kesilapan manusia serta ketidakkonsistenan hasil. Seiring perkembangan pesat dalam teknologi kecerdasan buatan (AI), pendekatan berasaskan pembelajaran mendalam telah diperkenalkan dalam bidang hematologi, khususnya untuk pengecaman, pengelasan, dan pengiraan sel darah putih secara automatik melalui analisis imej calitan darah.

Penggunaan sistem automatik berasaskan AI berpotensi meningkatkan kecekapan dan ketepatan dalam proses diagnostik serta mengurangkan kebergantungan kepada kaedah manual yang memakan masa. Walau bagaimanapun, terdapat beberapa cabaran yang masih wujud, terutamanya dalam menangani variasi kompleks dalam imej calitan darah, termasuk perbezaan morfologi sel, kualiti imej yang berubah-ubah dan objek bertindih yang boleh menjejaskan prestasi pengecaman (Zhao et al., 2020; Liang et al., 2022). Kajian terdahulu menunjukkan bahawa model seperti YOLO, meskipun mempunyai kelebihan dari segi kelajuan, kurang berkesan dalam mengesan objek bersaiz kecil, manakala model seperti Faster R-CNN menunjukkan ketepatan lebih tinggi tetapi dengan prestasi masa nyata yang terhad (Yao et al., 2021; Ghosh et al., 2021).

Tambahan pula, proses pengiraan automatik sel darah putih juga berdepan dengan cabaran ketepatan, terutamanya apabila sel-sel berada dalam kedudukan bertindih atau tidak jelas secara visual. Ketidaktepatan dalam pengiraan boleh mengakibatkan interpretasi diagnostik yang salah, yang seterusnya memberi kesan terhadap keberkesanan rawatan (Nugraha et al., 2023). Justeru, masih terdapat keperluan untuk satu sistem pintar berasaskan pembelajaran mendalam yang mampu melaksanakan pengecaman, pengelasan dan pengiraan sel darah putih dengan prestasi tinggi secara serentak.

Kajian ini dilaksanakan dengan matlamat untuk membangunkan sistem pengecaman, pengelasan dan pengiraan sel darah putih secara automatik menggunakan pendekatan pembelajaran mendalam. Pembangunan sistem ini diharapkan dapat menyumbang kepada peningkatan ketepatan diagnostik dalam bidang hematologi, khususnya dalam persekitaran klinikal yang memerlukan analisis yang pantas dan konsisten. Skop kajian ini merangkumi analisis imej calitan darah mikroskopik yang diperoleh dalam keadaan terkawal. Kajian ini menumpukan kepada penggunaan model pembelajaran mendalam untuk mengenal pasti dan mengelas jenis sel darah putih serta mengira bilangannya secara automatik. Kajian ini tidak melibatkan spesimen darah dalam bentuk lain seperti sampel cecair atau biopsi.

Bagi memastikan pembangunan sistem yang berkesan dan terkawal, pendekatan pembangunan spiral telah digunakan. Model spiral menggabungkan elemen prototaip dan kawalan sistematik dengan menekankan penilaian berterusan serta pengurusan risiko sepanjang kitaran pembangunan. Proses pembangunan dibahagikan kepada beberapa fasa, termasuk fasa perancangan, analisis risiko, pembangunan dan pengujian, serta penilaian. Setiap kitaran spiral membolehkan penambahbaikan dilakukan secara berulang berdasarkan maklum balas dan hasil eksperimen, sekaligus meningkatkan kualiti serta kebolehpercayaan sistem yang dibangunkan.

2.0 KAJIAN LITERATUR

Kajian literatur ini membincangkan pendekatan terdahulu dalam pembangunan sistem pengecaman, pengelasan, dan pengiraan sel darah putih secara automatik menggunakan pembelajaran mendalam. Kaedah tradisional yang bergantung pada pemeriksaan mikroskopik manual oleh pakar adalah kurang efisien serta terdedah kepada kesilapan penilaian dan keletihan manusia (Lee et al., 2022). Dengan kemajuan dalam bidang visi komputer, pelbagai model seperti YOLO dan DenseNet telah digunakan bagi meningkatkan kecekapan dan ketepatan dalam analisis imej perubatan (Zhao et al., 2020; Yao et al., 2021).

Model YOLO terkenal dalam tugas pengecaman objek masa nyata kerana kepantasannya, manakala DenseNet-121 pula digunakan untuk tugas pengelasan kerana keupayaannya mengekstrak ciri morfologi kompleks sel darah putih. Walau bagaimanapun, cabaran seperti variasi bentuk sel, kualiti imej yang tidak seragam, serta kehadiran objek bertindih terus menjejaskan ketepatan model (Nugraha et al., 2023). Oleh itu, kajian terdahulu telah mencadangkan penggunaan teknik pra-pemprosesan imej, augmentasi data dan penalaan hiperparameter sebagai usaha untuk menambah baik prestasi model.

2.1 PENGECAMAN DAN PENGELASAN SEL DARAH PUTIH MENGGUNAKAN PEMBELAJARAN MENDALAM

Kajian dalam bidang pengecaman dan pengelasan sel darah putih secara automatik semakin berkembang seiring dengan kemajuan teknologi pembelajaran mendalam dan pengimejan digital. Sistem tradisional yang bergantung kepada pemerhatian mikroskopik manual oleh pakar sering kali mengambil masa dan berisiko mengalami ralat manusia. Oleh itu, pendekatan automatik menggunakan kecerdasan buatan seperti Rangkaian Neural Konvolusi (CNN) dan model You Only Look Once (YOLO) telah diterokai secara meluas untuk meningkatkan kecekapan dan ketepatan proses ini.

Model CNN digunakan untuk mengekstrak ciri imej mikroskopik dan mengklasifikasikan jenis sel darah putih secara automatik. Antara model yang popular termasuk DenseNet-121, ResNet-50 dan EfficientNet-B0. Kajian menunjukkan bahawa DenseNet-121 boleh mencapai ketepatan setinggi 94%,

manakala EfficientNet-B0 pula mampu mencapai 97.37% dengan kelebihan dalam pemprosesan masa nyata. Sementara itu, model YOLO seperti YOLOv8 dan YOLOv9 digunakan untuk pengecaman serentak dan pengesanan sel dalam imej, yang penting untuk pengiraan dan klasifikasi dalam satu proses pantas dan efisien. Penggabungan YOLO dengan teknik moden seperti Transformer DETR turut menunjukkan peningkatan ketepatan yang ketara.

Pemilihan model bergantung kepada keperluan aplikasi: model YOLO sesuai untuk analisis pantas dan masa nyata, manakala CNN dua peringkat seperti Faster R-CNN lebih sesuai untuk situasi yang memerlukan ketelitian tinggi walaupun dengan masa pemprosesan yang lebih lama. Oleh itu, kombinasi antara ketepatan dan kecekapan pemprosesan menjadi asas penting dalam membangunkan sistem pengecaman dan pengelasan sel darah putih yang berpotensi digunakan secara klinikal.

2.2 PENGIRAAN SEL DARAH PUTIH MENGGUNAKAN TEKNIK PEMBELAJARAN MENDALAM DAN PRA-PEMROSESAN IMEJ

Pengiraan automatik sel darah putih amat penting dalam diagnosis penyakit hematologi. Namun, cabaran utama adalah apabila sel bertindih atau bentuknya berbeza-beza. Pelbagai pendekatan pembelajaran mendalam seperti YOLO, U-Net, dan Vision Transformer (ViT) telah digunakan bagi meningkatkan ketepatan pengiraan. Algoritma pengecaman objek seperti R-CNN, SSD dan YOLO mengenal pasti serta mengira sel berdasarkan bounding box. Kajian seperti oleh Wang et al. (2021) dan Lee et al. (2020) menggabungkan model seperti Fast Mask R-CNN dan VGG-16 untuk meningkatkan prestasi kiraan sel walaupun dalam situasi kompleks.

Selain pengecaman, pendekatan anggaran ketumpatan juga digunakan untuk mengira jumlah sel berdasarkan peta ketumpatan imej. Ini membantu dalam kes di mana sel bertindih atau tiada nukleus yang jelas. Contohnya, rangka kerja yang dicadangkan oleh Xue et al. (2021) dan He et al. (2020) menggunakan CNN dan FRCNN untuk memetakan imej ke peta ketumpatan yang tepat.

Proses pra-pemprosesan seperti penukaran imej ke skala kelabu, penyingkiran hingar, dan penormalan imej turut memainkan peranan penting. Ia memastikan imej yang digunakan lebih bersih, seragam dan sesuai untuk input model seperti ResNet dan DenseNet. Langkah ini terbukti meningkatkan prestasi model dalam banyak kajian.

Pemilihan dataset yang berkualiti turut penting. Beberapa dataset terkenal seperti BCCD, LISC, ALL-IDB, dan Free-Rabin menyediakan data untuk tugas klasifikasi, segmentasi dan pengecaman sel darah putih. Dataset ini berbeza dari segi bilangan kelas, jenis sel, dan resolusi imej. Justeru pemilihan dataset perlu selari dengan objektif kajian dan model yang dibangunkan.

3.0 METODOLOGI

Model pembangunan yang digunakan dalam kajian ini ialah **Model Spiral**, iaitu pendekatan pembangunan perisian berulang (iteratif) yang menggabungkan kelebihan daripada model prototaip dan model jujukan linear. Model ini dipilih kerana keupayaannya untuk menyokong proses pembangunan secara berperingkat, membenarkan penilaian berterusan, pengurusan risiko, serta pengubahsuaian berdasarkan maklum balas pengguna. Rajah 3.1 menunjukkan gambaran umum model pembangunan Spiral.



Rajah 3.1 Model Pembangunan Spiral

Sumber : <https://www.bdtask.com/blog/spiral-model-for-software-development>

Model ini amat sesuai digunakan dalam pembangunan sistem pengecaman, pengelasan dan pengiraan sel darah putih secara automatik kerana ia melibatkan proses pemodelan yang kompleks, pemprosesan imej digital dan interaksi pengguna yang memerlukan semakan berulang terhadap fungsi dan prestasi sistem. Setiap kitaran pembangunan sistem ini merangkumi empat fasa utama iaitu fasa perancangan, fasa analisis risiko, fasa pembangunan dan pengujian dan fasa penilaian.

3.1 Fasa Perancangan

Fasa perancangan memberi tumpuan kepada penentuan skop sistem dan keperluan pengguna. Keperluan ini didokumentasikan dalam Spesifikasi Keperluan Sistem (SRS) yang merangkumi keperluan fungsional seperti muat naik imej, pengecaman sel, klasifikasi automatik, dan pengiraan bilangan sel, serta keperluan bukan fungsional seperti prestasi, ketepatan, dan kebolehgunaan sistem.

3.2 Fasa Analisis Risiko

Dalam fasa ini, risiko-risiko teknikal seperti kegagalan pengecaman objek, prestasi model yang rendah, kesesuaian data, dan kebolehgunaan sistem dikenalpasti. Proses ini termasuk pembangunan prototaip awal sistem pengecaman menggunakan model YOLOv8 dan pengelasan menggunakan DenseNet-121

untuk menilai potensi prestasi model pembelajaran mendalam yang digunakan. Jika risiko utama dikenalpasti, penyesuaian terhadap model, pra-pemprosesan data, atau seni bina sistem akan dicadangkan.

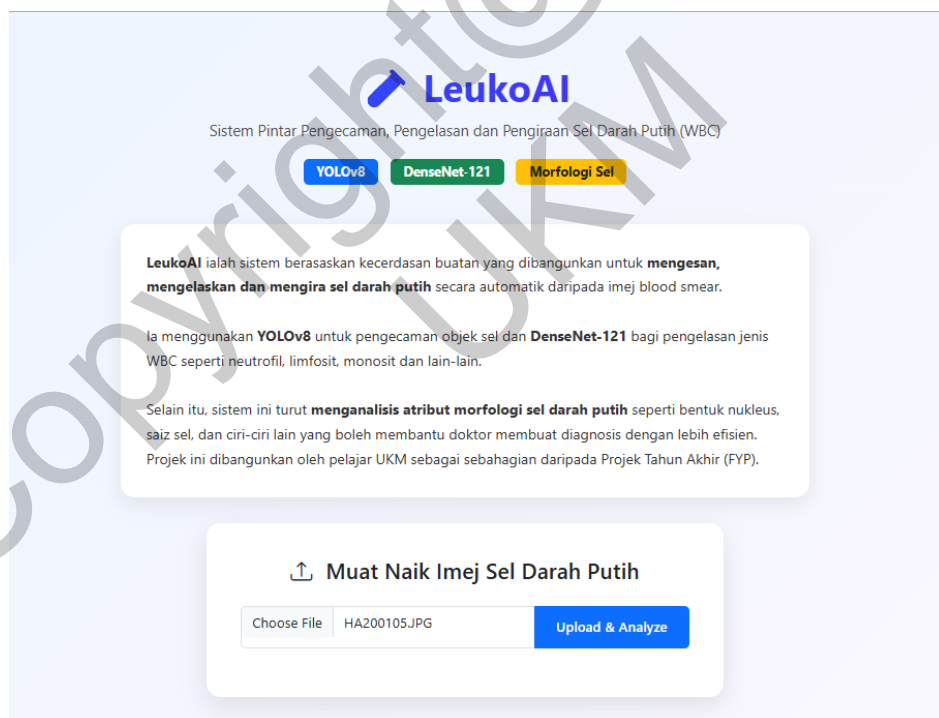
3.3 Fasa Pembangunan dan Pengujian

Sistem dibangunkan secara modular berdasarkan reka bentuk seni bina berlapis seperti yang diuraikan dalam seksyen 3.5. Model YOLOv8 digunakan untuk pengesanan sel darah putih manakala DenseNet-121 digunakan untuk proses pengelasan. Pembangunan melibatkan antara muka pengguna, logik pemprosesan imej, dan integrasi model AI.

Antara muka pengguna yang dibangunkan terdiri daripada:

i. Antara muka muat naik data:

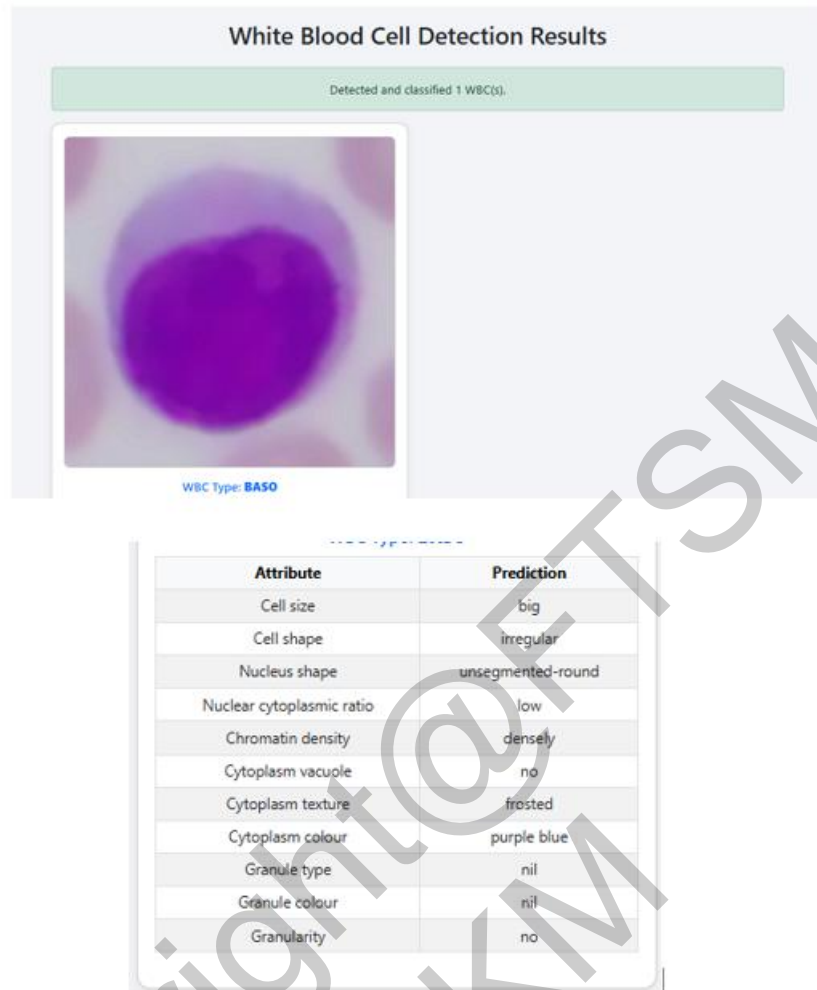
Pengguna memuat naik imej calitan darah ke dalam sistem untuk diproses oleh model pengesanan dan pengelasan.



Rajah 3.2 Rajah Antara Muka Muat Naik Data

ii. Antara muka paparan keputusan:

Sistem memaparkan output dalam bentuk jadual hasil pengesanan, klasifikasi, dan pengiraan sel darah putih bagi tujuan pemahaman dan sokongan diagnosis oleh pengguna.



Rajah 3.3 Rajah Hasil Keputusan Data

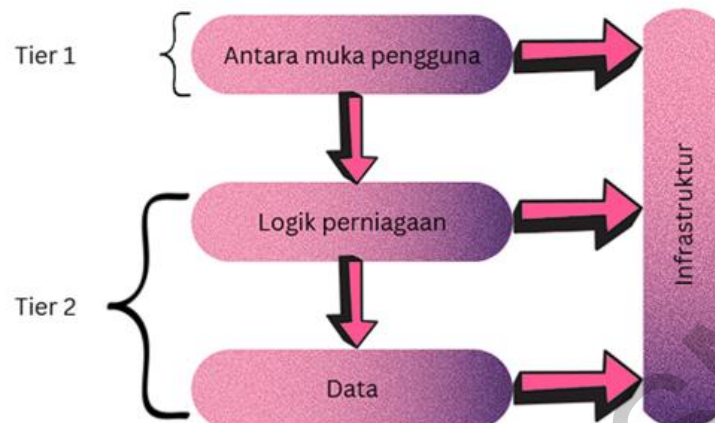
Pengujian dilaksanakan secara berterusan bagi setiap modul melalui ujian fungsi dan ujian ketepatan model. Fasa ini turut menyokong pembangunan algoritma sistem dan reka bentuk antara muka pengguna.

3.4 Fasa Penilaian

Output daripada setiap iterasi pembangunan akan dinilai bersama pengguna sasaran seperti pelajar perubatan atau penyelidik. Maklum balas diperoleh melalui ujian penggunaan sistem secara simulasi dan digunakan untuk penambahbaikan reka bentuk serta fungsi sistem dalam iterasi seterusnya. Kriteria penilaian termasuk ketepatan pengecaman sel, kebolehgunaan sistem, dan masa pemprosesan.

3.5 Reka Bentuk Seni Beni Berlapis

Sistem pengecaman, pengelasan dan pengiraan sel darah putih ini dibangunkan menggunakan reka bentuk seni bina berlapis yang terdiri daripada Tier 1 (Antara Muka Pengguna) dan Tier 2 (Logik Perniagaan dan Data). Reka bentuk ini membolehkan sistem dibina secara modular, mudah diselenggara dan lebih fleksibel.



Rajah 3.4 Rajah Seni Bina Berlapis

Tier 1 menyediakan antara muka yang membolehkan pengguna memuat naik imej smear darah, menjalankan analisis dan melihat hasil pengecaman serta pengelasan sel darah putih secara visual. Tier 2 mengandungi komponen utama sistem. Model YOLOv8 digunakan untuk mengecam dan melokalisasi sel darah putih dalam imej mikroskopik, manakala DenseNet-121 digunakan untuk mengklasifikasikan sel kepada jenis seperti neutrofil, limfosit, dan lain-lain. Semua data imej, label, dan hasil analisis diuruskan dalam lapisan data sistem.

Sistem ini turut disokong oleh infrastruktur berasaskan GPU bagi memastikan pemprosesan imej dilakukan dengan pantas dan cekap. Pendekatan berlapis ini menyokong skalabiliti dan kebolehselenggaraan sistem dalam aplikasi perubatan.

4.0 HASIL KEPUTUSAN

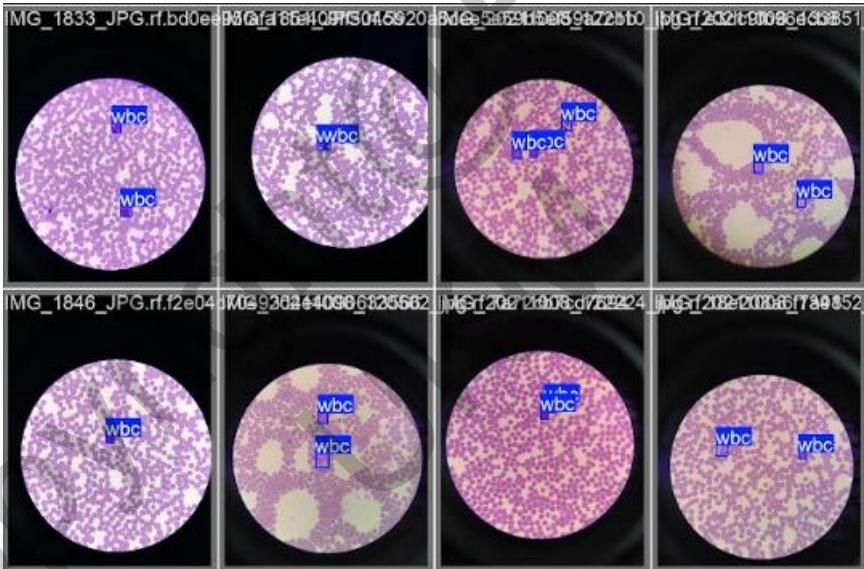
Bahagian ini membincangkan hasil prestasi model yang digunakan dalam sistem pengecaman, pengelasan dan pengiraan sel darah putih. Dua model utama telah diuji iaitu YOLOv8 untuk pengecaman sel darah putih dan DenseNet-121 untuk pengelasan jenis sel. Penilaian dilakukan berdasarkan metrik standard seperti ketepatan, kadar pengesanan semula, skor F1 dan matriks kekeliruan.

4.1 Model YOLOv8

Model YOLOv8 digunakan untuk mengecam sel darah putih dalam imej mikroskopik. Latihan dilakukan menggunakan dataset yang telah dianotasi, dengan input imej bersaiz 640×640 piksel selama 50 epoch. Model ini menunjukkan prestasi pengecaman yang sangat tinggi.

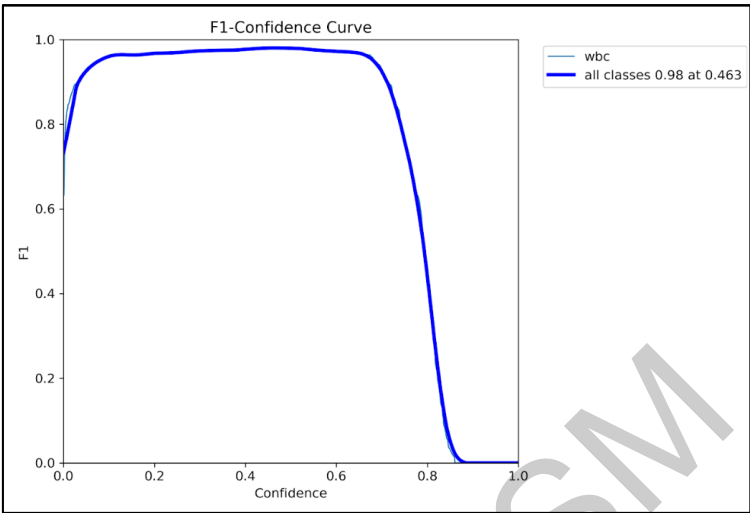
Jadual 4.1 Jadual Keputusan Pengecaman

Metrik Penilaian	Nilai
Ketepatan Positif (Precision)	0.992
Kadar Pengesanan Semula (Recall)	0.969
mAP@0.5	0.988
mAP@0.5:0.95	0.671



Rajah 4.1 Rajah Imej Pengecaman validasi

Menunjukkan bahawa setiap sel darah putih berjaya dikesan dan dilabel dengan tahap keyakinan 0.6–0.9.



Rajah 4.2 Graf Keyakinan Skor F1

Graf ini menunjukkan hubungan antara nilai keyakinan (confidence) dan skor F1. Skor F1 merupakan purata harmonik antara precision dan recall. Model ini mencapai skor F1 tertinggi 0.98 pada ambang keyakinan 0.463, menandakan keseimbangan optimum antara ketepatan dan keupayaan mengesan objek. Prestasi model kekal konsisten ($F1 > 0.95$) dalam julat keyakinan antara 0.2 hingga 0.75, membuktikan kestabilan model dalam pelbagai senario.



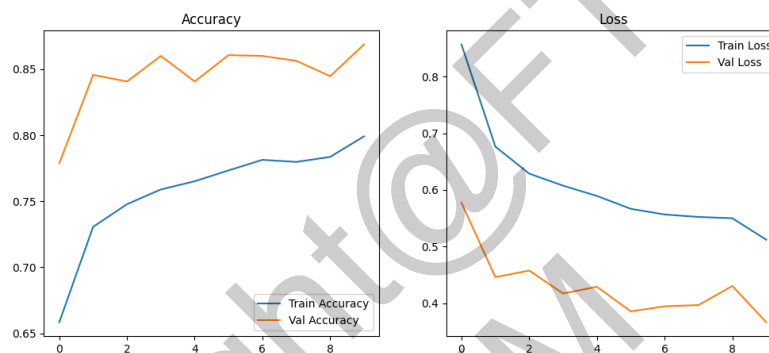
Rajah 4.3 Matriks Kekeliruan

Matriks ini menunjukkan pengedaran prediksi berbanding kelas sebenar. Model berjaya mengenal pasti 355 imej wbc dengan betul (*True Positive*), hanya 10 kes gagal dikesan (*False Negative*) dan 11 *False Positive*. Matriks ini menunjukkan ketepatan pengesanan melebihi 96%, dengan kadar kesilapan yang sangat rendah.

4.2 Model DenseNet-121

Model DenseNet-121 telah digunakan dalam sistem ini bagi melaksanakan tugas klasifikasi lima jenis utama sel darah putih iaitu basofil, monosit, neutrofil, limfosit dan eosinofil. Pemilihan model ini adalah berdasarkan keupayaannya membina sambungan langsung antara lapisan-lapisan terdahulu, yang membantu mengekalkan dan menyebarkan maklumat ciri dengan lebih efisien. Struktur padat model ini juga sangat sesuai untuk pengenalan imej perubatan yang memerlukan pengekstrakan ciri kompleks pada resolusi tinggi.

Graf dalam Rajah 4.4 menunjukkan perkembangan prestasi model sepanjang sesi latihan selama 10 epoch. Sub-graf sebelah kiri memaparkan ketepatan bagi data latihan dan penilaian, manakala sebelah kanan menunjukkan kerugian.



Rajah 4.4 Graf Ketepatan(kiri) dan Graf Kerugian(Kanan)

Terdapat peningkatan konsisten dalam ketepatan latihan, daripada 66% kepada 80%, menunjukkan bahawa model mampu belajar dengan berkesan dari set latihan. Keepatan penilaian pula mengekalkan prestasi stabil pada sekitar 84% hingga 87% selepas epoch ke-1, menandakan tiada masalah overfitting yang ketara. Manakala, latihan kerugian menurun secara berterusan daripada 0.86 kepada 0.52, menunjukkan pengurangan ralat dalam ramalan model. Penilaian kerugian juga menurun dan kekal rendah sekitar 0.37 ke 0.45, yang mengesahkan kestabilan model dalam menangani data baharu.

Model telah diuji menggunakan set data ujian yang terdiri daripada 3,000 imej sel darah putih, yang mengandungi 600 imej bagi setiap lima kelas utama. Hasil ujian menunjukkan bahawa model DenseNet-121 mencapai ketepatan keseluruhan sebanyak 97.67%, yang menggambarkan keupayaan model untuk melaksanakan tugas klasifikasi dengan sangat tinggi dan tepat.

Penilaian lanjut menggunakan metrik ketepatan positif, kadar pengesanan semula, dan skor F1 menunjukkan bahawa model mengekalkan prestasi yang konsisten dan seimbang bagi semua kelas. Kesemua lima kelas mencatatkan nilai skor F1 melebihi 0.96, dengan kelas basofil memperoleh nilai tertinggi iaitu 1.00, manakala kelas lain seperti eosinofil, limfosit, monosit dan neutrofil mencatatkan

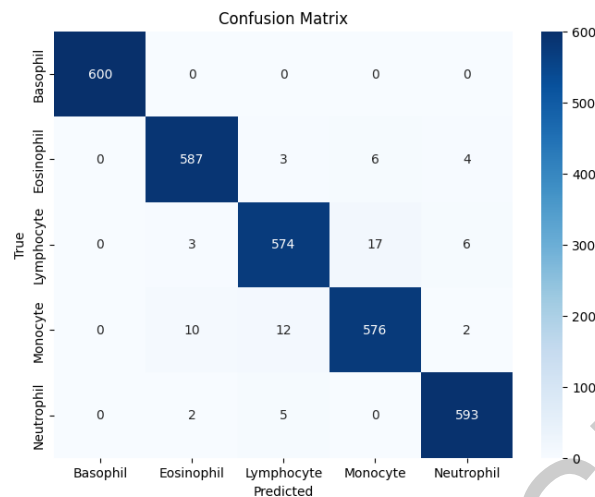
skor yang sangat hampir kepada sempurna. Hal ini membuktikan keberkesanan model dalam mengekalkan keseimbangan antara kadar pengesanan semula dan ketepatan positif.

Prestasi terperinci bagi setiap kelas ditunjukkan dalam Jadual 4.2:

Jadual 4.2 Jadual Keputusan Pengelasan

Kelas	Precision	Recall	F1-Score
Basophil	1.00	1.00	1.00
Eosinophil	0.98	0.98	0.98
Lymphocyte	0.97	0.96	0.96
Monocyte	0.96	0.96	0.96
Neutrophil	0.98	0.99	0.98

Selain itu, Rajah 4.5 menunjukkan matriks kekeliruan bagi model apabila diuji dengan set data ujian. Berdasarkan matriks ini, dapat dilihat bahawa setiap kelas diklasifikasikan dengan sangat baik, sebagai contoh, model mengklasifikasikan kesemua 600 imej basofil dengan tepat. Bagi kelas lain seperti neutrofil, sebanyak 593 daripada 600 imej diklasifikasikan dengan betul. Jumlah kesilapan adalah kecil dan tidak signifikan, contohnya kelas limfosit tersalah klasifikasi sebagai monosit sebanyak 17 imej sahaja, yang merupakan ralat minor dalam konteks keseluruhan prestasi.



Rajah 4.5 Matriks Kekeliruan Model Pengelasan

Secara keseluruhannya, model DenseNet-121 menunjukkan prestasi cemerlang dalam klasifikasi sel darah putih dengan keputusan yang konsisten tinggi dalam semua metrik penilaian. Ini membuktikan bahawa model bukan sahaja mampu belajar dengan berkesan semasa latihan, malah dapat mengaplikasikan pengetahuannya secara kukuh pada data baharu. Justeru, model ini berpotensi untuk digunakan dalam sistem sokongan diagnosis imej perubatan sebenar secara automatik dan dipercayai.

5.0 KESIMPULAN

Projek LeukoAI telah berjaya dibangunkan sebagai satu sistem pintar yang menggabungkan teknologi pembelajaran mendalam bagi melaksanakan pengecaman, pengelasan dan pengiraan sel darah putih secara automatik berdasarkan imej mikroskopik calitan darah. Sistem ini menggunakan model YOLOv8 untuk pengecaman objek dan DenseNet-121 untuk klasifikasi imej, yang kemudiannya diintegrasikan ke dalam antara muka web menggunakan rangka kerja Flask. Hasil ujian menunjukkan bahawa sistem ini mampu memberikan output yang tepat, mesra pengguna dan dapat menangani ralat input secara efektif. Secara keseluruhan, LeukoAI memberi sumbangan bermakna dalam bidang pemprosesan imej perubatan dan analisis hematologi dengan menyediakan satu platform automatik yang efisien untuk tujuan penyelidikan dan pendidikan.

Walau bagaimanapun, terdapat beberapa kekangan seperti ketidakseimbangan data bagi kelas minoriti seperti basofil, serta ketidakstabilan model klasifikasi apabila berdepan imej kabur atau tidak konsisten dari segi pencahayaan. Sistem ini juga masih kekurangan fungsi simpanan data dan validasi pakar, menjadikannya belum sesuai untuk aplikasi klinikal sebenar. Bagi penambahbaikan, pengumpulan data berskala besar, integrasi model segmentasi yang lebih tepat, pembangunan sistem login pengguna serta sokongan pengesahan keputusan oleh pakar boleh dipertimbangkan. Pembangunan versi aplikasi mudah alih juga boleh memperluas capaian dan impak sistem ini dalam konteks penggunaan dunia sebenar.

RUJUKAN

- Ahmad, R., Awais, M., Kausar, N., & Akram, T. (2023, January 18). White blood cells classification using entropy-controlled deep features optimization. MDPI. <https://www.mdpi.com/2075-4418/13/3/352>
- Alam, M. M., & Islam, M. T. (2019, July 17). Machine learning approach of automatic identification and counting of blood cells. IET Research Journals. <https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/htl.2018.5098>
- Almezhghwi, K., & Serte, S. (2020, July 9). Improved classification of white blood cells with the generative adversarial network and deep convolutional neural network. Wiley Online Library. <https://onlinelibrary.wiley.com/doi/full/10.1155/2020/6490479>
- Bendibdallah, M. H., & Settout, N. (2023, February 14). Deep learning for segmentation and counting of white blood cells in clinical diagnosis. ResearchGate. https://www.researchgate.net/publication/370857589_Deep_learning_for_segmentation_and_counting_of_white_blood_cells_in_clinical_diagnosis
- Cleveland Clinic. (n.d.). White blood cells. Cleveland Clinic. <https://my.clevelandclinic.org/health/body/21871-white-blood-cells>
- El-Seoud, S. A., Siala, M., & McKee, G. (2020, December). Detection and classification of white blood cells through deep learning techniques. ResearchGate. https://www.researchgate.net/publication/347883408_Detection_and_Classification_of_White_Blood_Cells_through_Deep_Learning_Techniques
- Ferdousi, J., Lincoln, S. I., Alom, M. K., & Foysal, M. (2024, December). A deep learning approach for white blood cells image generation and classification using SRGAN and VGG19. *Telematics and Informatics Reports*, 16, 100163. <https://www.sciencedirect.com/science/article/pii/S2772503024000495>

- Higuera, V. (2024, March 7). What to know about white blood cell (WBC) count. Healthline. <https://www.healthline.com/health/wbc-count>
- Lee, S. J., Chen, P. Y., & Lin, J. W. (2022, August 14). Complete blood cell detection and counting based on deep neural networks. MDPI. <https://www.mdpi.com/2076-3417/12/16/8140>
- Nugraha, S. J. A., & Erfianto, B. (2023, July 31). White blood cell detection using Yolov8 integration with DETR to improve accuracy. ResearchGate. https://www.researchgate.net/publication/372930158_White_Blood_Cell_Detection_Using_Yolov8_Integration_with_DETR_to_Improve_Accuracy
- Pamungkas, A. S., Soelistio, Y., & Budiawan, R. (2024). White blood cell classification using EfficientNet for malaria diagnosis. Scientific Reports. (Disunting mengikut kandungan asal – sila sahkan butiran jika diperlukan).
- Sharma, A., Kumar, V., & Longchamps, L. (2024). Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and Faster R-CNN models for detection of multiple weed species. ResearchGate. <https://www.researchgate.net/publication/385691124>
- Sultana, K., Ahmed, H. T., Umair, M., Iftikhar, A., & Sultana, K. (2023, February 7). Covid-19 variants detection & classification using self-proposed two-stage MNN-2: Robust comparison with Yolo V5 & Faster R-CNN. IEEE. <https://ieeexplore.ieee.org/abstract/document/10034649>
- Tamang, T., Baral, S., & Paing, M. P. (2022, November 22). Classification of white blood cells: A comprehensive study using transfer learning based on convolutional neural networks. PMC. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9777002/>
- Toğaçar, M., Ergen, B., & Cömert, Z. (2020, December). Classification of white blood cells using deep features obtained from convolutional neural network models based on the

combination of feature selection methods. *Applied Soft Computing*, 97, 106810.

<https://www.sciencedirect.com/science/article/abs/pii/S1568494620307481>

Veeraiah, N., Alotaibi, Y., & Subahi, A. F. (2022, December 21). MayGAN: Mayfly optimization with generative adversarial network-based deep learning method to classify leukemia from blood smear images. *Computer Systems Science & Engineering*. <https://doi.org/10.32604/csse.2023.036985>

Worsito, S. B., & Pumrin, S. (2022, January 6–7). Image-based white blood cell subtype detection using ResNet-50 convolutional neural network (CNN). *RC-EEE 2021 Proceedings*. https://www.researchgate.net/profile/DMandal/publication/365098606_rc-eee2021-proceeding-2022M3-18/links/637711c11766b34c5433bf5d/rc-eee2021-proceeding-2022M3-18.pdf#page=171

Xu, Z., Zou, J., Long, D., Wei, X., & Yang, G. (2024). Blood cell detection: Enhanced YOLOv9 algorithm. *IEEE*. <https://ieeexplore.ieee.org/abstract/document/10865573>

Yao, J., Huang, X., Wei, M., Han, W., Xu, X., Wang, R., Chen, J., & Sun, L. (2021, September 13). High-efficiency classification of white blood cells based on object detection. *Wiley Online Library*. <https://onlinelibrary.wiley.com/doi/10.1155/2021/1615192>

Nur Umairah Nabilah Binti Khairudiin (A196052)

Ts. Dr. Afzan Adam

Fakulti Teknologi & Sains Maklumat

Universiti Kebangsaan Malaysia