

PENDETEKSIAN SEL DARAH PUTIH DARI CITRA PREPARAT DENGAN CONVOLUTIONAL NEURAL NETWORK

Danny ¹⁾ Lina ²⁾ Arlends Chris ³⁾

¹⁾²⁾³⁾ Teknik Informatika Universitas Tarumanagara

Jl. Letjen S. Parman No. 1, Jakarta 11440 Indonesia

¹⁾dindratjuatja@gmail.com ²⁾lina@untar.ac.id ³⁾arlendsc@fk.untar.ac.id

Abstract

White Blood Cells play an important role as parts of the immune system by fighting against viruses, bacteria and potentially harmful foreign objects that enter the human body. The amount of white blood cells can indicate a certain disease or infection within the human body. This research aims to develop a system that can automatically detect and locate the location of white blood cells in a slide image that is stained or not stained. By not staining blood cell images, it can save time and resources that are normally used in white blood cell detection. This system is built using convolutional neural networks (CNN), a deep learning architecture. The CNN model is used for detecting white blood cells in stained images and is trained with 528 images and the model that is used for detecting white blood cells in unstained images is trained with 264 images. Bounding box regression is used to predict the location of white blood cells. The experiment test results show the detection accuracy for the stained images reach 53.85% and for the unstained images has 54.69% accuracy.

Key words

Convolutional Neural Network, Deep learning, Detection, HSV, White Blood Cells

1. Pendahuluan

Pada tubuh manusia terdapat sel darah merah, sel darah putih dan keping-keping darah. Sel darah merah berfungsi untuk mengantar oksigen dari paru-paru ke seluruh tubuh dan mengangkut karbondioksida dari jaringan tubuh kembali ke paru-paru. Sel darah putih merupakan salah satu bagian dari sistem imun manusia yang melindungi tubuh dari berbagai bakteri, virus dan objek asing lainnya dengan cara membunuh mereka [1]. Sel darah putih dapat dibagi menjadi lima yaitu, neutrofil, basofil, eosinofil, limfosit dan monosit.

Pemeriksaan darah merupakan salah satu cara pemeriksaan medis untuk mendiagnosa berbagai penyakit dan juga untuk mengetahui golongan darah [2]. Pada darah terdapat sel darah putih, sel darah merah dan keping-keping darah. Jumlah sel darah putih yang tinggi ataupun rendah dapat menunjukkan adanya

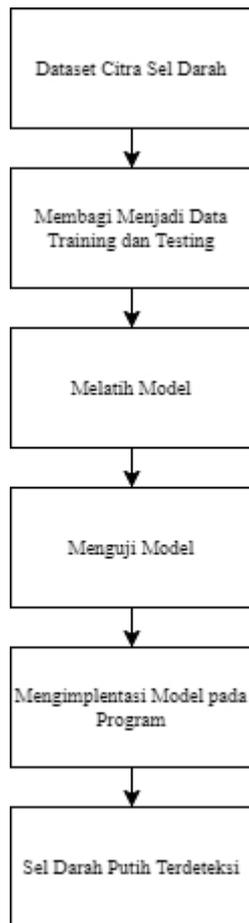
penyakit tertentu atau infeksi tertentu yang terjadi pada tubuh. Jika jumlah sel darah putih berdasar di atas jumlah rata-rata maka dapat menunjukkan penyakit kanker darah sedangkan jumlah sel darah putih yang berada di bawah jumlah seharusnya dapat mengindikasikan leukopenia yang dapat disebabkan oleh infeksi berat.

Dalam pemeriksaan darah, jumlah sel darah putih pada sebuah volume darah dapat memberi informasi kondisi pasien. Proses pendeteksian sel darah putih yang dilakukan saat ini adalah dengan cara inspeksi visual oleh ahli medis dibawah mikroskop dengan membutuhkan larutan kimia yaitu, larutan Wright dan Buffer dalam pemberian warna pada citra preparat sel darah. Larutan tersebut berguna agar sel darah putih terlihat lebih kontras pada mikroskop sehingga memudahkan dalam pendeteksian sel darah putih.

Dengan berkembangnya teknologi, maka dapat dimanfaatkan dalam membantu pendeteksian sel darah putih secara otomatis dan dapat mendeteksi sel darah putih pada citra preparat dengan pewarnaan dan tanpa pewarnaan untuk membantu para ahli medis dalam melakukan pemeriksaan darah. Oleh karena itu, dilakukan penelitian untuk mendeteksi sel darah putih pada citra preparat dengan pewarnaan dan tanpa pewarnaan. Namun hasil pendeteksian yang dilakukan pada rancangan sebelumnya masih memiliki tingkat keakuratan kurang dari 50% pada citra preparat tanpa pewarnaan [3]. Pada percobaan ini, maka dikembangkan sistem yang dapat mendeteksi sel darah putih dengan pewarnaan dan tanpa pewarnaan serta meningkatkan tingkat keakuratan sebelumnya.

2. Metode Penelitian

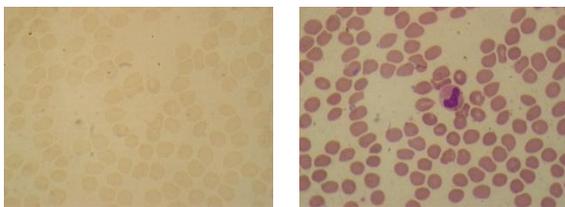
Pada percobaan ini, sistem dibangun dengan menggunakan *Convolutional Neural Network* (CNN), salah satu arsitektur dari *deep learning*. Sistem dibangun untuk mendeteksi sel darah putih pada citra preparat dengan pewarnaan dan tanpa pewarnaan. Model yang dihasilkan akan diimplementasikan pada program. Tahapan pada sistem secara umum dapat dilihat pada Gambar 1.



Gambar 1. Skema Sistem

2.1. Dataset

Pada percobaan ini, dataset yang digunakan merupakan data citra sel darah yang diambil pada preparat sel darah. Data citra sel darah diambil menggunakan kamera yang disambungkan ke mikroskop. Perbesaran yang digunakan pada mikroskop adalah perbesaran 1000 kali, didapat dengan perbesaran 10 kali pada lensa okuler dan 100 kali pada perbesaran lensa objektif. Contoh data citra preparat sel darah dengan pewarnaan dan tanpa pewarnaan dapat dilihat pada Gambar 2.



Gambar 2. Citra Sel Darah Tanpa Pewarnaan (Kiri) dan Citra Sel Darah Dengan Pewarnaan (Kanan)

Dataset yang didapat merupakan citra sel darah dengan pewarnaan (*stained*) dan citra sel darah tanpa pewarnaan (*unstained*). Setiap dataset akan dibagi menjadi data latih yang digunakan untuk melatih model dan data uji untuk menguji hasil model. Data yang

berhasil dikumpulkan berjumlah 1050 citra. Tabel jumlah dataset yang digunakan pada percobaan ini dapat dilihat pada Tabel 1.

Tabel 1. Tabel Data

Tipe	Jumlah
WBC (<i>stained</i>)	361
No WBC (<i>stained</i>)	361
WBC (<i>unstained</i>)	164
No Wbc (<i>unstained</i>)	164
Total	1050

2.2. Convolutional Neural Network

Convolutional neural network merupakan kategori dari Artificial Neural Network (ANN), yang sudah terbukti bagus dalam klasifikasi citra [4]. merupakan metode *deep learning* yang terdiri dari berbagai lapisan tersembunyi. Lapisan tersebut dapat dibagi menjadi dua, yaitu lapisan ekstraksi fitur untuk mendapatkan karakteristik dari masukan dan layer klasifikasi untuk menentukan kelas [5]. Dalam mendapatkan lokasi citra sel darah putih pada citra maka digunakan *bounding box regression* yang akan dilatih dengan menggunakan citra yang sudah diberi label posisi sel darah putih. Lapisan *feature map* yang digunakan dalam pengklasifikasian citra juga digunakan dalam *bounding box regression*.

Pelatihan model dilakukan dengan epoch sebanyak 30 kali dan dengan *batch size* 64. Pada percobaan ini, terdapat dua jenis dataset citra preparat dengan pewarnaan dan tanpa pewarnaan. Pelatihan dilakukan secara terpisah pertama dilakukan pelatihan untuk citra preparat dengan pewarnaan. Pada citra preparat dengan pewarnaan maka digunakan *pre-trained* model VGG16 yang didalamnya terdapat 16 lapisan konvolusi yang digunakan. Konfigurasi layer yang digunakan pada pelatihan citra preparat dengan pewarnaan dapat dilihat pada Tabel 2.

Tabel 2. Konfigurasi Lapisan Citra Preparat Darah Dengan Pewarnaan

Lapisan	Parameter
VGG16	
Conv2D	64, (3,3), Relu
Conv2D	32, (3,3), Relu
Dropout	0.2
Flatten	-

Dense	2, Sigmoid
-------	------------

Pelatihan model untuk citra preparat tanpa pewarnaan juga menggunakan epoch sebesar 30 kali dan *batch size* sebesar 64. Pada pelatihan citra preparat tanpa pewarnaan digunakan *pre-trained* model ResNet50 yang memiliki 50 lapisan di dalamnya. Konfigurasi lapisan untuk citra preparat tanpa pewarnaan dapat dilihat pada Tabel 3.

Tabel 3. Konfigurasi Lapisan Citra Preparat Tanpa Pewarnaan

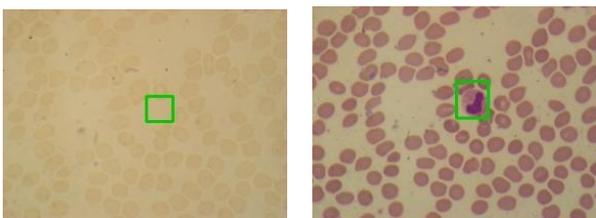
Lapisan	Parameter
ResNet50	
Conv2D	64, (3,3), Relu
Conv2D	32, (3,3), Relu
Dropout	0.2
Flatten	-
Dense	2, Sigmoid

Setelah latihan selesai, didapatkan hasil pelatihan. Hasil pelatihan dari model yang dilatih dapat dilihat pada Tabel 4.

Dataset	Loss	Accuracy	Val_Loss	Val_Accuracy
Preparat Darah (Stained)	0.2315	93.45%	0.5346	82.18%
Preparat Darah (Unstained)	0.3952	80.58%	0.4568	71.27%

3. Hasil Pengujian

Pengujian dilakukan dengan cara melihat hasil deteksi lokasi sel darah putih model. Model yang telah dilatih, akan dilakukan pengujian untuk mengevaluasi hasil dari proses pelatihan. Contoh pendeteksian dengan hasil yang benar oleh model dapat dilihat pada Gambar 3.

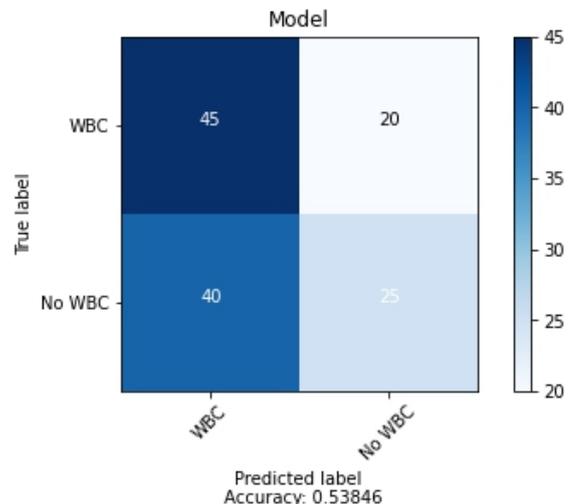


Gambar 3. Contoh Pendeteksian Benar

Pengujian dilakukan dengan menggunakan dataset uji yang berbeda dari dataset pelatihan untuk mengetahui tingkat keakuratan dari model yang telah dilatih pada citra yang berbeda dengan citra pelatihan. Pengujian menggunakan data sebanyak 130 citra, 65 citra terdapat sel darah putih dan 65 citra tidak terdapat sel darah putih. Metode pengujian menggunakan *confusion matrix*. *Confusion Matrix* adalah metode yang digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada *confusion matrix* terdapat nilai *true positive*, *false positive*, *true negative* dan *false negative*. Keterangan untuk setiap klasifikasi data hasil pengujian adalah sebagai berikut:

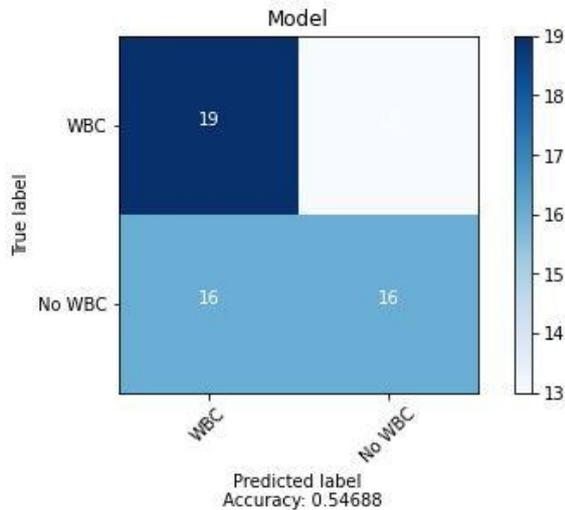
1. True Positive = Citra dengan WBC, terdeteksi
2. False Positive = Citra tanpa WBC, terdeteksi
3. True Negative = Citra tanpa WBC, tidak terdeteksi
4. False Negative = Citra dengan WBC, tidak terdeteksi

Hasil pengujian model yang digunakan untuk citra preparat pewarnaan yang berupa *confusion matrix* dapat dilihat pada Gambar 4.



Gambar 4. Confusion Matrix Model Citra Preparat Dengan Pewarnaan

Hasil pengujian model berupa *confusion matrix* yang digunakan untuk pendeteksian citra preparat tanpa pewarnaan dapat dilihat pada Gambar 5.



Gambar 5. Confusion Matrix Model Citra Preparat Tanpa Pewarnaan

Dari confusion matrix yang telah diperoleh, maka dapat dihitung juga nilai akurasi, presisi, recall dan F1-Score. Nilai akurasi dapat dihitung dengan menggunakan rumus sebagai berikut:

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

Nilai presisi dapat dihitung dengan menggunakan rumus sebagai berikut:

$$Presisi = \frac{TP}{TP+FP} \times 100\% \quad (2)$$

Nilai recall dapat dihitung dengan menggunakan rumus sebagai berikut:

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

Nilai F1-score dapat diperoleh dengan menggunakan rumus sebagai berikut:

$$F1\ Score = 2 \times \frac{Presisi \times Recall}{Presisi+Recall} \times 100\% \quad (4)$$

Dengan menggunakan rumus-rumus tersebut maka didapat performa dari model yang telah dibuat. Hasil dari performa model dalam nilai akurasi, presisi, recall dan F1-score dapat dilihat pada Tabel 5.

Skenario	Akurasi	Presisi	Recall	F1-Score
Model Preparat Darah (Stained)	53.25%	54.17%	69.23%	60%
Model Preparat Darah (Unstained)	54.69%	54.29%	59.38%	56.72%

Hasil pengujian dengan menggunakan data yang berbeda menghasilkan akurasi yang tidak sebagus dari hasil saat pelatihan. Dengan melihat hasil ini, dapat dibilang bahwa model yang digunakan mungkin terjadi *overfitting*, dimana model terlalu mempelajari pola yang ada pada data latih tetapi tidak terdapat pada data uji.

4. Kesimpulan

Kesimpulan yang diperoleh dari percobaan yang telah dilakukan terhadap pendeteksian sel darah putih dengan CNN adalah sebagai berikut:

- Hasil pengujian terhadap pendeteksian sel darah putih pada citra preparat sel darah dengan pewarnaan memiliki akurasi 53.25%. Sedangkan hasil pengujian pendeteksian sel darah putih pada citra preparat tanpa pewarnaan memiliki tingkat akurasi 54.69%. Pengujian dilakukan dengan menggunakan data sebanyak 130 data untuk setiap modelnya.
- Hasil pelatihan memiliki angka akurasi mencapai 93.45% pada citra preparat pewarnaan dan 80.58% untuk citra preparat tanpa pewarnaan, tetapi pada pengujian hasil yang diperoleh kurang baik. Hal ini dapat disebabkan oleh model yang *overfitting*.

Dalam pengembangan untuk kedepannya, dapat menggunakan epoch yang lebih banyak atau konfigurasi lapisan yang berbeda ataupun *pre-trained* model yang berbeda untuk memperoleh hasil yang lebih baik lagi dan tidak overfit. Serta melakukan optimasi *preprocessing* pada citra sel darah tanpa pewarnaan agar ciri karakteristik sel darah putih dapat terlihat lebih jelas lagi.

REFERENSI

- [1] Wang, Yapin dan Cao, Yiping, 2019, "Leukocyte Nucleus Segmentation Method Based On Enchancing The Saliency of Saturation Component", Journal of Algorithm and Computational Technolgy, Vol. 13.
- [2] Habibzadeh, Mehdi, Krzyzak, Adam, dan Fevens Thomas, 2013, "White Blood Cell Differential Counts Using Convolutional Neural Networks for Low Resolution Images", ICAISC 2013, Part II. Berlin.
- [3] Reynaldo, David, 2019, "Pendeteksian Sel Darah Putih Pada Citra Preparat Tanpa Pewarnaan Dengan Watershed Transformation", Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara,
- [4] Tiwari et al, 2018, "Detection of Subtype Blood Cells Using Deep Learning", Cognitive System Research, Vol. 52.
- [5] Zufar, Muhammad dan Setiyono, Budi, 2016, "Convolutional Neural Networks untuk Pengenalan Wajah Secara Real-Time", Jurnal Sains dan Seni ITS, Vol. 5, No. 2. Surabaya: Institut Teknologi Sepuluh.

Danny, saat ini sebagai mahasiswa Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara.

Lina S.T., M.Kom., Ph.D., memperoleh gelar Sarjana dari Universitas Tarumanagara, Indonesia tahun 2001 dan gelar Magister dari Universitas Indonesia, Indonesia tahun 2004. Kemudian tahun 2009 memperoleh gelar Ph.D. dari Nagoya University, Jepang. Saat ini sebagai Dosen Tetap Program Studi Teknik Informatika Fakultas Teknologi Informasi Universitas Tarumanagara.

Arlends Chris S.Ked., dr., M.Si., Dr., memperoleh gelar Sarjana dari Universitas Tarumanagara, Indonesia dan gelar Magister dari Universitas Tarumanagara, Indonesia. Kemudian memperoleh gelar Ph.D. pada tahun 2019 dari Universitas Negeri Jakarta, Indonesia. Saat ini sebagai Dosen Tetap Fakultas Kedokteran Universitas Tarumanagara.